Essays on Mutual Fund Investor Attention

by

Michael John Wymbs

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2021 by the
Graduate Supervisory Committee:

George Aragon, Chair
Yuri Tserlukevich
Oliver Boguth

ARIZONA STATE UNIVERSITY

May 2021

ABSTRACT

I propose new measures of investor attention for Mutual Funds. Using the Security and Exchange Commissions' Electronic Data Gathering, Analysis, and Retrieval (EDGAR) system's server log files, this study is the first to explore investor attention to specific mutual funds. I find that changes, or spikes, in mutual fund investor attention are associated with funds' introduction of a new share class, decreases in expense ratio, past performance and volatility. On average, spikes to investor attention predict net inflows into mutual funds which outpace the overall growth of the mutual fund sector. Attention via this EDGAR channel is more important when investors are researching more opaque funds. Moreover, there is a positive relationship between mutual fund investor attention and fund returns. Yet, there is evidence that investors appear to be responding to the acquisition of stale information with flows. I additionally utilize Google Trends data for individual fund tickers and investigate its effects in Mutual Fund Market. I find that Investor Attention to individual mutual funds is concentrated within Equity funds, Index funds, and Institutional funds. Individual fund attention is strongly negatively associated with expense ratios, 12B-1 Fees, and 'broker sold' funds, suggesting that funds with higher fees get less attention than low cost index funds. I find limited support for the controversial convexity in the flow to performance sensitivity in the Mutual Fund market, but only in funds with high levels of individual attention.

DEDICATION

To Mom, Dad, Julia and Kate.

ACKNOWLEDGEMENTS

I am grateful for my advisors, George Aragon, Yuri Tserlukevich and Oliver Boguth for their guidance on this work. I especially would like to thank my chair, George Aragon, for his positive encouragement and insightful ideas that fostered the development of this doctoral thesis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

1) INTRODUCTION

There is a large empirical literature studying mutual fund flows due to their impact on fund performance, asset prices, fund manager compensation and incentives. The mutual fund sector is a massive[1] savings and investment vehicle for households[2]. Fund flows give insight into the time-varying preferences of the investing public. In response to flows, the delegated portfolio managers make trading decisions for underlying securities which can affect security prices and overall fund performance. Moreover, these managers' compensation is directly tied to levels of assets under management, creating incentives for managers to attract as much flows as possible. Mutual fund characteristics, such as prior returns, rating, and fees, are associated with predicting flows, suggesting that investors care about some observable fund characteristics when forming their asset allocation decisions. Other types of information, such as those contained in regulatory filings, could be relevant for investors. However, no one has researched if the attention of mutual fund investors to specific funds can predict flows and its consequences.

Why does investor attention matter? Any investor who desires to invest in a mutual fund must have paid attention to the fund before any transaction occurs, as Merton (1987) postulates. Investors do not invest in investments of which they are not ex-ante aware. This paper gains insight into what specific funds are selected to be screened by potential investors. I construct a novel measure attention of mutual fund investors through web traffic of the SEC's EDGAR website of users examining mutual fund regulatory filings, contributing to the fund flow prediction literature. The direct mechanism assumes investors are utilizing EDGAR to perform due diligence on their investment opportunity set. The examination of a filing is a strong signal that a purchase could be imminent, as there are more funds to be invested in than are currently held for a given investor.

[1] Mutual funds held $18.7 trillion of US assets at the end of 2017, including almost a quarter of U.S. public equity

[2] According to the ICI 2019 factbook, households held 89% of mutual fund assets at year-end 2018.

From the sever log files of the Security and Exchange Commission's EDGAR system, I can determine the precise times investors, uniquely identifiable throughout the sample with a partially masked IP address, accessed a specific fund's filing. I aggregate views of filings across funds for a given time period and calculate the share of attention received by a fund relative to the entire viewed universe of EDGAR. I then construct my measure, Abnormal View Share, as spikes in the time series of share of views for a given fund as deviations from a trailing six month moving median. It is intended to capture the dynamics of attention consistent with the existing literature on Investor Attention (e.g. Da, Engelberg and Gao (2011), Ben-Rephael, Da and Israelsen (2017)). Determinants of Abnormal View Share can be explained by several factors, such as fund age, assets under management, additional fund creation, and lowering of expenses, but not long term returns and funds ratings.

As mutual funds cannot be short sold, information acquisition in this setting should be signed positively. I find, indeed, that Abnormal View Share of mutual fund investor attention positively predict net flows into mutual funds. These flows out pace the overall growth of the mutual fund sector. For example, a one standard deviation increase in Abnormal View Share is associated with a significant incremental increases in CIK-level net flows of 2.3 basis points per month, after controlling for known fund flow predictors. Abnormal View Share is positively associated with subsequent new inflows into funds but insignificant with respect to outflows.

What information are investors looking at within the filings that is aiding their decision making process? From views of forms 485BPOS and 497, I examine what information content within viewed filings investors care about, using fund-level flows as evidence. I find that excess returns over the stated benchmark and management fees in viewed filings predict flows to funds after subsequent to view of the funds' filing. Additionally, I find evidence that stale information in filings which are over a year old also predicts flows after controlling for known predictors.

Grossman and Stiglitz (1980) argue that because obtaining and processing information is costly, in equilibrium, those who acquire must be doing so because they believe the benefits of the information acquisition outweigh the costs, that is "they can use their information to take positions in the market which are 'better' than the positions of uninformed traders". Do the selected investments for which investors acquire information outperform? I find abnormal mutual fund investor attention is positively related to future fund performance. A long short decile portfolio sorted on Abnormal View Share earns a CAPM Alpha of 17.44 basis points per month. Sorting decile portfolios on an Abnormal View Share measure counting only filings viewed within the month of release to the public has even better performance, suggesting that those investors who keep up with fund regulatory news select funds to view that ex-post outperform. However, most of the predictability comes from the short portfolio, meaning sudden inattention of fund filings is a leading signal of future under-performance. This result potentially survives arbitrage because mutual fund cannot be directly sold short.

The Abnormal View Share calculated may not actually be measuring Mutual Fund Investors' attention, rather it may instead be a proxy of global sentiment for bullishness in US Markets. In order to distinguish, I can partition views from domestic and international based on the IP Addresses in the sample. International viewers are assumed to have more difficult access to investments in US based mutual funds than US based mutual fund investors, and as a result, may not be reflecting "Investor" attention. I find that domestic attention is positively associated with future fund flows and fund returns, while foreign spikes in attention are unrelated to both flows and returns.

The acquisition of information from this regulatory filing channel appears to benefit investors, because it aids in their asset allocation decision making process. This is especially true for funds that are ex-ante more opaque, as I document these less well known funds have a strong EDGAR attention-flow relationship. However, the views of regulatory filings it could also be potentially harmful to investors, as mandated disclosure of certain fund information, such as portfolio holdings, could leave mutual funds vulnerable to potential "copycat" competitors. Cao, Du, Yang, and Zhang (2019) find evidence that some 13-F Filers, upon viewing their competitor's 13-F filings, subsequently trade in the same direction, mitigating potential returns. I find that views coming from IP Addresses that belong to financial institutions do not predict flows, but are positively associated with subsequent returns of the viewed funds, suggesting there may be other reasons that fund filings get viewed other than for evaluation of investment opportunities.

The rest of the paper is organized as follows. Section 2 describes the related literature. Section 3 discusses the data set and preparation. Section 4 investigates the determinants of Investor Attention. Section 5 relates attention to flows. Section 6 investigates abnormal investor attention and returns. Section 7 concludes.

2) LITERATURE

Measuring demand of investor attention for mutual funds has unique challenges relative to proxies for attention which are typically used in equity related studies. Funds do not have trading volume (Barber and Odean (2008)) like equities which are directly bought and sold. Google searches for fund tickers (Da, Engelberg and Gao (2011)) are very sparse, most fund-month observations do not have the minimum amount of searches for them to be measurable this way, in part because funds have many share classes each with unique tickers. News viewership on Bloomberg terminals (Ben-Rephael, Da and Israelsen (2017)) is also very sparse for fund specific news relative to equity specific news, as there are many more 'newsworthy events' for individual companies than portfolios.

While funds do not have periodic earnings announcements, both mutual funds and equities are companies regulated by the SEC, and as such are required to file registration statements, periodic reports, and other forms electronically through EDGAR.[3] Regulatory filings, in essence, are a form of fund specific news which can be accessed by investors on the SEC's EDGAR[4] system. Investors obtain and view these filings electronically, acquiring information about funds to inform their investment decisions. The view of a particular fund's filing directly and unambiguously implies attention has been paid to the fund.

This paper is the first to use the EDGAR log file dataset to measure demand of mutual fund investor attention. This is in contrast to previous literature which relates proxies to the supply side of investor attention, e.g. news about a fund (Sirri and Tufano (1998)), mutual fund advertising (Jain and Wu (2002)) or components of fund holdings (Solomon, Soltes, and Sosyura (2014)), to equity mutual fund flows. EDGAR Log Files have been examined previously by

[3] https://www.sec.gov/edgar.shtml

[4] Electronic Data Gathering, Analysis, and Retrieval

researchers, with a substantial focus driven toward equity filings. First, Bauguess, Cooney and Hanley (2018) investigate the investor demand for newly issued equity securities. Drake, Roulston and Thornock (2014) investigate the determinants of information acquisition for equity's filings. They find investors focus on a subset of filings, leaving the vast majority of filings rarely requested. Loughran and McDonald (2014), (2017) examine when filings are viewed, finding that 10-K filings are not accessed by many investors immediately after they are released, suggesting that investors are seeking out information from EDGAR for reasons other than trades on filing dates. Lee, Ma, and Wang (2015) apply a co-search algorithm to identify peer related firms. Gibbons, Iliev, and Kalodimos (2019) show equity analysts utilize EDGAR filings, reducing forecasting errors. Dyer (2017) finds institutional investors acquire more information, more quickly about local stocks. Li and Sun (2017), Wang (2019), and Drake, Johnson, Roulstone and Thornock (2019), relate EDGAR views of firms to future firm performance. Chen, Cohen, Gurun, Lou and Malloy (2017) observe that institutional investors track insider trade filings and form profitable trading strategies. Other work demonstrating the prowess of sophisticated investors utilizing EDGAR filings include Crane, Crotty, and Umar (2019) which demonstrates that hedge fund investors profit from attention to equities, while Chen, Kelly and Wu (2019) show that hedge fund investors in response to brokerage closures increase efforts to view affected equity filings and earn higher returns in a Grossman and Stiglitz (1980) framework. Iliev, Kalodimos and Lowry (2019) examine the monitoring of mutual fund viewers on their equity holding's proxy statements filed to EDGAR. Cao, Du, Yang, and Zhang (2019) examine 13-F filers in EDGAR who view other 13-F filings and find evidence of copycatting behavior.

Considerably less work has been performed with respect to investor attention within the mutual fund sector, a gap in the literature this paper hopes to fill. The two related papers to the idea of mutual fund investor attention are Sicherman, Loewenstein, Seppi and Utkus (2016) which studies investor attention paid to personal Vanguard retirement accounts, which are predominantly invested in "Vanguard stock, bond, balanced, and money market mutual funds" and Kim (2017) which studies mutual fund families creating Twitter accounts to capture investor attention and attract flows, but cannot distinguish attention to different funds within the family.

3) DATA

Following a Freedom of Information Act (FOIA) request by the public, the SEC has made the server logs of EDGAR publicly available. The log files provide daily requests to EDGAR from Jan 1, 2003-June 30, 2017.[5] My sample consists of all filings associated with the SEC's Central Index Key (CIK) that can be mapped to CRSP Mutual Fund Universe. The unit of observation at the filing level is the CIK, which is a required logon information for EDGAR for any filer to the SEC, individuals or companies. New filing entities can gain access to a CIK by filling out Form-ID[6]. While the form explicitly states "This application is for potential filers who are new to the SEC only: if the potential filer (i.e., Applicant) already has an assigned EDGAR Central Index Key (CIK), do not use this form! In this case, refer to Volume I, Chapter 3.3 of the EDGAR Filer Manual for more information.", there has been a lack of standardization regarding how mutual fund management companies have chosen to satisfy their filing requirements with respect to grouping their mutual funds within CIKs. For most mutual fund management companies in CRSP, the CRSP MGMT_CD maps one to one with the CIK (1015 maps), however some management companies have grouped funds together under separate trusts, e.g., the Fidelity Concord Street Trust contains distinct funds and a different CIK than the Fidelity Aberdeen Street Trust. To complicate things further, some management companies such as American Funds or Oppenheimerfunds have separate CIK entries for different funds. Ultimately, the filer has discretion in its choices to set up multiple CIKs or not, in order to satisfy the SEC's disclosure requirements. Table 1 Panel A displays the 10 largest CIKs by assets, and the number of share classes associated with funds within these CIKs. Note that Vanguard for example chooses to group different sets of funds into 5 different CIKs within the top 10.

The log files contain the filing requested, time and date of the request, the associated CIK of the filing, and the IP Address displayed in IPv4 format, with the 4th number masked, e.g.

[5] April 29-30, 2017 are missing from the dataset.

[6] located https://www.filermanagement.edgarfiling.sec.gov/filermgmt/selectFormId.html

199.67.131.jag. The masking was intended to preserve the uniqueness of the IP address throughout the sample. The displayed numbers of the IP address are enough information to identify physical locations of the IP address.[7] Similarly, institutions will typically buy up large blocks within the IP Address universe. With the use of MaxMind GeoIP linking file, the block of "199.67.131" can be identified as belonging to Citigroup, which contains all IP Addresses between 199.67.128.0 and 199.67.138.255, e.g. 199.67.128.0 - 199.67.128.255,199.67.127.0 - 199.67.127.255, ... 199.67.138.0 - 199.67.138.255. I match these organizations names to 297 financial organizations that contain at least a hedge fund in TASS among their subsidiaries. These also correspond to 62 organizations that contain at least one mutual fund family among their subsidiaries. I dichotomize those IP Addresses that come from the 297 financial organizations as Financial from those that do not, which I will refer to as Retail. I also remove the views from mutual fund families who are viewing their own filings.

CRSP provides a linking table to SEC's CIKs that maps 90% of the CRSP universe by assets. I hand collect missing CIK's from the mutual fund's name in CRSP to obtain 99.5% coverage by TNA.[8] I exclude those CIKs that contain a 10-K or 10-Q filing, as these corporate disclosures can bias results pertaining to mutual fund investor attention. I require observations to have TNA data in CRSP. I value weight all characteristics of funds within the CIK. I obtain Morningstar ratings for fund-months during the sample period from Morningstar Direct, and merge them to CRSP fundnos through CUSIPs. I value weight fund ratings[9] to the CIK level.

I remove 'robot' generated downloads according to Ryans (2017). Specifically, I classify an IP Address-Day pair as robotic if one of the following criteria are met: 1) The log file itself classifies the user as a "crawler". 2) The IP address downloads 500 or more documents on a given day. 3) The IP address downloads 25 or more documents in a given minute. 4) The IP

---

[7] I reference https://db-ip.com/db/download/ip-to-city-lite to obtain information up the city level. When there are IP blocks that are split e.g. in the 4th number 0-200 correspond to city1 and 201-255 correspond to city2, I assume the address maps to city1, the most probable match. In identifying countries, I am 99.9% confident the country is mapped correctly.

[8] The majority of unmatched consists of electronically traded notes, which do not have the same filing requirements.

[9] Where available, as funds less than 3 years old do not have a rating

address downloads documents from 3 or more different CIKs within a minute. I remove duplicate filing views defined as the same IP address viewing the same filing on a given day. Additionally, I filter views of identifiable fund management company viewers viewing their own filings.

Figure 1 plots the monthly views of mutual fund filings after filtering. As the data is sparse in the early part of the sample, I do all subsequent analysis from 01/2007-06/2017. There is an exponential growth rate of views beginning in 2011. Other papers (e.g. Gibbons, Iliev, and Kalodimos (2019), Wang (2019)) use log number of filing views as their main measure of interest, but mechanically this leads to larger observations in the later part of the sample period. As time has progressed, technological progress has made web based content more accessible, yielding lower search costs of EDGAR filings for investors. Additionally, over time more investors have become aware of EDGAR as a means for accessing SEC filings. Moreover, the universe of available filings to view is expanding daily. To control for this time trend in filing views I aggregate the filtered EDGAR filing views to the monthly frequency for each CIK, and then scale by all mutual fund views in a given month. I define EDGAR View Share for a given month t, and CIK i as:

$$EDGARViewShare_{i,t} = {FilingsViewed_{i,t}}\Big/{\sum_{i=1}^{n} FilingsViewed_{i,t}}$$

While the caveats outlined by Bauguess, Cooney and Hanley (2018) apply with use of the EDGAR log file dataset[10], the entire log file sample represents the universe of filing views through this channel, e.g., coming from the SEC's website, during the sample period. My EDGAR View Share measure allows me to compare the fraction of views between two time periods, given ex-ante any filing is equally accessible on the site. Table 1 Panel B displays the CIKs with largest time series average view share during the sample period. The filings associated with the "iShares Trust" and "PIMCO Funds" CIKs receive more than 17x the share of views than the average CIK in the sample.

[10] SEC filings can be obtained outside of EDGAR, from Bloomberg for example. Filings are also obtainable from the SEC's ftp site which are not included in this sample. EDGAR log files under-represent actual demand for filing information.

The EDGAR View Share measure and its log(Filing Views) counterpart are measures of the *levels* of investor attention. However, the Investor Attention literature (e.g. Da, Engelberg Gao (2011); Ben-Rephael, Da, and Israelsen (2017)) has been more interested in *spikes* to the time series of levels of attention rather than attention in levels itself, to capture the dynamics of investors' information acquisition.[11] As such, my main variable of interest is Abnormal View Share in the time series defined as the spike in EDGAR View Share relative to the median EDGAR View Share for the previous six months. For a given month t, and CIK i:

$$AbnormalViewShare_{i,t} = EDGARViewShare_{i,t} - Median\left(EDGARViewShare_{i,t-1}, \ldots, EDGARViewShare_{i,t-6}\right)$$

Table 1 Panel C ranks the top 10 CIKs by largest Abnormal View Share on average throughout the sample. Panel D ranks the top 10 filing types by number of views in the sample. Table 2 displays the summary statistics of the Mutual Fund CIK level characteristics. For each characteristic, I compute the value weighted measure of the corresponding characteristic for each component funds within the CIK.

4) DETERMINANTS OF INVESTOR ATTENTION OF EDGAR FILINGS FOR MUTUAL FUNDS

What are the determinants of Abnormal View Share? Why do some funds at certain times receive a larger share of the views than in previous times? In Table 3, lagged fund characteristics are regressed on Abnormal View Share. Overall, the majority of the variation of Abnormal View Share remains unexplained.

The most significant predictor of a spike in the share of attention of Mutual Fund filings is a dummy variable for an increase in the number of share classes associated with the CIK. The certeris paribus effect of one or more new share classes being added to a CIK is the equivalent of

[11] Both Google Trends based measures of search interest and Bloomberg based measures of news readership are identifying spikes in the time series relative to the baseline level attention of a given security.

an 11% standard deviation increase in Abnormal View Share[12]. Investors pay attention to filings when new fund options are added to the CIK's menu. Decreases in share classes associated with the CIK also garner attention spikes but only 38% of the effect of increasing share classes. Larger CIKs in terms of assets under management are more likely to receive additional attention. A one standard deviation increase in the log total net assets for a CIK is associated with an 26% increase in a funds Abnormal View Share attention relative to the mean AVS level[13]. Analogous to larger firms being more attention grabbing, it is unsurprising that CIKs that encompass a greater number of share classes garner more interest because CIK size is not standardized, e.g., if there are more funds within the group then there are more chances for an individual investor seeking out information for a given fund to look within that group. Monthly returns, but not annual returns, predict spikes in attention. A one standard deviation increase in the past month return of the CIK is associated with 15% increase in attention. Further, investors appear to have asymmetric responses to changes in expense ratio, a one standard deviation decrease in the value weighted level of the expense ratio of the CIK is associated with a 4.6% increase in the share of attention by investors, but no significant effect in changes in view shares is observed for net increases in fund expenses. CIKs that are on average older receive less attention through the EDGAR channel, as more newly created funds have associated filings that garner interest. Abnormal attention is also associated with less volatile funds, funds with larger levels of expenses, and ETFs. Other factors such as Morningstar ratings or index fund composition do not offer significant explanatory power in this model.

5) ABNORMAL EDGAR ATTENTION AND MF CIK LEVEL FLOWS

While in principle attention itself is a neutral outcome, e.g. both good news and bad news precipitate attention[14], Barber and Odean (2008) claim that attention is (1) a driver of purchases

[12] The standard deviation of Abnormal View Share is 4.74 basis points.
[13] The mean Abnormal View Share is 0.35 basis points.
[14] Prior work in Investor Attention relates attention measures neutral outcomes such as trading volume or absolute value of returns or price change.

but not sales of individual equities, and (2) does not apply with equal force to institutional investors. They reason when buying, retail (institutional) investors ex-ante have small (large) portfolios relative to space of available securities, whereas when selling, retail investors typically only sell what they own, while institutional investors sell securities short with a greater frequency than retail investors. Unlike equities, mutual funds cannot be directly sold short. This institutional detail implies that positive fund signals obtained through fund information acquisition are more likely to result in changes in market positions than negative fund signals. Moreover, there may be motivations unrelated to information acquisition such as liquidity needs or tax consequences that may be driving sell decisions.

Investor attention, ex-ante, may not matter at all with respect to mutual fund investment decision making. Investors could just be chasing past returns when making their decisions regarding where to invest capital, as is well documented in the flow performance literature, e.g. Sirri and Tufano (1998). Morningstar ratings of funds have been shown to have a significant incremental impact on investor flows (Del Guercio and Tkac (2008); Evans and Sun (2020)). It is possible that variation in fund flows that is correlated with observed investor attention is explained by prior fund performance and/or Morningstar rating. I find that while investor attention is associated with aggregate net fund flow increases on average, different populations of funds may result in a non-relation with investor attention.

5.1) Abnormal Attention and Flows – Aggregate Analysis

I obtain net flows monthly for a given CIK from aggregating the change in assets in excess of value weighted returns of funds within an individual CIK group, correcting for inter-CIK fund mergers and liquidated funds[15]. Given that a large channel for Mutual Fund flows are through retirement savings channels via defined contribution or defined benefit plans, I attempt to control for this effect by including the previous month's Net CIK Flow percentage in the set of dependent variables. I use the following regression specification:

[15] To control for outliers in the sample, I winsorize flows at the 2% level.

$$NetCIKFlow(\%)_{i,t} = \alpha + \beta * AbnormalViewShare_{i,t} + \gamma * MFControls \ + \ \delta_t + \varepsilon_{i,t}$$

In Table 4, I find in the full specification of the CRSP US fund universe, Abnormal View Share is significantly positively related to future flows. Holding other factors constant, a one standard deviation increase in Abnormal View Share is associated with a significant incremental increases in CIK-level net flows by 2.3 basis points per month.[16] Past returns, both annual and monthly, and Morningstar ratings are strong positive predictors of mutual fund flows. CIKs containing ETFs are positively associated with additional flows. The valued weighted expense ratio of the CIK appear to be an insignificant predictor of subsequent flows. Investors respond to declines in expense ratio with subsequent flows. Other characteristics behave as one might expect, flows are inversely related to fund size (Berk and Green (2004)), fund volatility, and fund age.

5.2) Abnormal Attention and Flows – Mechanism

The econometrician observing net flows cannot distinguish if attention is driving net buying activities or if inattention precedes net selling activities. To gain insight to test Barber and Odean (2008)'s claim in the mutual fund setting, I utilize data obtained from N-SAR filings of funds within the sample period. Form N-SAR was a semi-annual report filed by all register investment companies which contains information, at the fund level, of new inflows and outflows of capital at the monthly frequency[17]. I aggregate these monthly new inflows and outflows of capital to the CIK level, and regress new inflows and outflows, scaled by assets, on Abnormal View Share and Mutual Fund characteristic controls. I find, in Table 5, that Abnormal View Share is positively associated with subsequent new inflows but insignificant with respect to outflows, which is consistent the Barber and Odean (2008)'s predictions for equities. Interestingly, the past year return is insignificantly related to new inflows, but significantly negatively related to outflows. This suggests that the mechanism regarding how investor attention is associated with net flows is different from past performance. Poor long term performance is more likely to be punished

[16] One standard deviation in Abnormal View Share = .000474; .4846 * .000474 = .000230
[17] New inflows here is capital distinguished from re-invested capital.

through outflows than consistent long term performance is to be rewarded through new inflows of capital.

For robustness, I test the dependent variable of Spiegel and Zhang (2013)'s Change in Market Share. This scales CIK asset growth by all assets in the U.S. Mutual Fund market (e.g. CRSP universe), thus I am able to test if spikes in the share of views of EDGAR filings predict growth in assets at the CIK level that out pace the growth of assets to the entire mutual fund sector. I find that after controlling for mutual fund characteristics, Abnormal View Share significantly predicts positive changes in CIK market share. Ceteris paribus, a one standard deviation increase in Abnormal View Share predicts an annualized increase in market share of .1 basis points for the CIK.[18] Increasing investor attention to a particular fund within the EDGAR filing channel is associated with incremental growth that outpaces the growth of the mutual fund industry in the United States.

5.3) Abnormal Attention and Flows – Opacity

Some mutual funds are more well known and transparent than others. Mutual fund attention presumably could matter more to the investor when funds are more opaque. To test if there is a greater attention-flow response for more opaque funds I use two proxies for opaqueness in my empirical design.

More opaque funds may not be rated by Morningstar. Morningstar requires a three year return history to assign a fund rating, but does not automatically assign a rating to every fund once they have 36 months of returns. However, once a fund becomes rated by Morningstar it is more easily searched for on their site, and is presumably more well known by investors given the documented importance of Morningstar ratings (Del Guercio and Tkac (2008); Evans and Sun (2020)). Ideally, I would be able to test attention to the specific fund before and after it receives a rating, but the unit of observation here is the CIK, consisting of some funds that are rated and some that are not, so I'm forced to use a cruder proxy. In Table 6 specification (1), I interact Abnormal View Share with an indicator variable, "Post Morningstar Rated Dummy", defined as all

[18] .000474*14.7061*12 = .084

CIK-months within the sample period when, as of the beginning of the sample period, the first not rated fund within the CIK obtains a rating. The coefficient on the interaction term is negative, significant, and in absolute value the same size as the coefficient on Abnormal View Share by itself, indicating that prior to obtaining a rating, there is a strong positive abnormal attention-flow response, that once a rating is obtained, is completely negated. When the CIK as a whole has become marginally less opaque, because a component fund within the CIK has received a rating, spikes in views from EDGAR do not predict flows. When investors cannot rely on a Morningstar rating to summarize information related to a fund EDGAR filings are a better source of information for potentially interested investors.

Within a similar line of reasoning, younger funds have less of a track record, are smaller in terms of assets, and less well known than older, established funds. I define a CIK as "Young" if the value weighted fund age is in the bottom quartile of the distribution[19]. I interact, in specification (2) a dummy variable for Young Funds with Abnormal View Share in the full specification regressing onto next month's Net CIK Flows. I find that Abnormal View Share, without an interaction is insignificant, and when interacted with a "young" CIK is positive, significant, and with a coefficient that is four times larger than in the same baseline model in Table 4.

In specification (3), I form a triple interaction model with Post Morningstar Rated and Young Fund dummies on Abnormal View Share. I find a positive significant coefficient on the interaction between Abnormal View Share and "Young", that is about ten times the size of the baseline model in Table 4. The triple interaction coefficient, as in the first specification is negative, significant and of the same magnitude in absolute value as the the interaction coefficient between Abnormal View Share and "Young", indicating that after the fund obtains a Morningstar rating, the net effect of Abnormal View Share of young funds on flows is negated. EDGAR appears to be more valuable of a resource to investors when they are researching more opaque funds.

[19] The cutoff point here is an average age of 90 months, or 7.5 years.

5.4) Abnormal Attention and Flows – Sub-samples

In order to better understand the effect of ratings on abnormal attention to mutual funds on mutual funds' flows, I perform the regressions in my full specification model in various sub-samples. I partition the space of CIKs into Active Equity Funds, Active Bond Funds, ETFs and Index Funds based on if the CIK has a majority of assets of the respective type.

Table 7 reports the coefficient on the Abnormal View Share variable across different sub-samples based on fund style and with partitioned measures of Abnormal View Share. I find little to no predictability for measures of Abnormal View Share on the Active Bond and Index Fund sub-samples, indicating that mutual fund investor attention with respect to flow prediction is concentrated in terms of Active Equity and ETFs.

First, I partition the views that get counted towards Abnormal View Share based on when the filing is viewed. If the filing is viewed within 30 days of being released to the public, I define this as "Recent" and conversely, in excess of 30 days I define as "Old". I find that the attention flow response is concentrated among the set of "Old" filing views for Active Equity funds and ETFs, and I see little evidence of spikes to attention of new filings predicting flows in these sub-samples, suggesting that old, archived information can still be valuable for investors in their due diligence process. There are some investors who keep current with new filings, but they do appear to respond to their views of these with subsequent flows.

The Abnormal View Share measure may not actually be reflecting Mutual Fund Investors' attention, instead it may be a proxy of global sentiment for bullishness in US Markets. In order to distinguish, I partition views from domestic and international based on the IP Addresses in the sample. International viewers are assumed to have more difficult access to investments in US based mutual funds than US based mutual fund investors, and as a result, may not be reflecting "Investor" attention. I find that domestic attention is positively associated with future fund flows, while foreign spikes in attention are unrelated to flows.

Finally, I examine if IP Addresses belonging to financial companies viewing mutual fund filings behave differently with respect to flows than the residual, which I label as "Retail". Does this crude partition of presumable investor sophistication result in different outcomes with respect to flows? I find that indeed, the retail views are associated with flows in the Active Equity and ETF sub-samples, while the spikes in financial views are unrelated to subsequent flows.

5.5) Information Content Observed within Filings

If investors are paying attention to filings in the investment decision making process, what information within the filing matters the most? To attempt to answer this question I have moved to a more narrow setting; "XBRL" filings, forms 497 and 485BPOS. The advantage of looking at XBRL filings is that not only are they the most viewed filings in the sample, but their information content has been tagged and structured by the SEC in the Mutual Fund Prospectus Risk/Return Summary Data Sets.

> "The Mutual Fund Prospectus Risk/Return Summary Data Sets provides text and numeric information extracted from the risk/return summary section of mutual fund prospectuses. The data is extracted from exhibits to mutual fund prospectuses tagged in eXtensible Business Reporting Language (XBRL). The information is presented without change from the "as filed" submissions by each registrant as of the date of the submission. The data is presented in a flattened format to help users analyze and compare corporate disclosure information over time and across registrants."[20]

This data set consists of forms 485BPOS and 497 beginning in January 2011. The other advantage of using the Mutual Fund Prospectus Risk/Return Summary Data Sets is that I am able to observe information on a more granular level than the fund's CIK. For each fund, the SEC assigns a Series ID and Class ID which can identify funds at the share class level. While the EDGAR log file data set only identifies the filing at the CIK level, the Mutual Fund Prospectus Risk/Return Summary Data Sets provide the Series ID and Class ID for individual data points within the filing, e.g. "Average Annual Return of Fund Since Inception". While some filings in the sample contain prospectuses for multiple funds, the extra level of identification will allow me to associate all share classes that are contained within the filing, and remove those that are not in

[20] Available https://www.sec.gov/dera/data/mutual-fund-prospectus-risk-return-summary-data-sets

the filing, but otherwise associated with the CIK. The unit of observation here is the filing-fund(s), value weighted if multiple exist.

I test if numeric information contained within a viewed filing is correlated with net flows of fund(s) in the month proceeding the view. The information tested is fees expressed as management fee percentage, 12b-1 fee percentage and net expenses percentage, Average annual return of fund in the past 1, 5, and 10 years and since inception, along with the average annual return of the stated, self-designated benchmark within the prospectus in the past 1, 5, and 10 years and since inception. Note, these numbers are only current as of when the filing is filed, not necessarily when the filing is ultimately viewed.

In Table 8, I find there is explanatory power in the information obtained by viewing a filing on subsequent month fund flows, after controlling for fund-level characteristics. This evidence is consistent with the Sensoy (2009)'s observation that self-designated benchmarks influence fund flows. He shows that performance relative to a specific benchmark of the fund is a significant determinant of net fund flows at the annual frequency. In my setting, I find the returns of the fund during the past year weigh positively, and returns of the self-designated benchmark in the past year weigh negatively on flows associated with the viewed fund in the month after it is viewed. Note that the returns observed are with respect to when the filing is filed, which is not necessarily in the same month or year that the filing is ultimately viewed by the investor. That is, the viewing of the historical out-performance of the fund with respect to the self-designated benchmark in the past year, from the filing date, is positively associated with net inflows subsequent to the information being acquired by the investor, consistent with Sensoy (2009).

However, I observe the opposite relation with respect to the returns of the fund and its benchmark when the returns are measured "Since Inception" of the fund. The observed Average Annual Return of the Fund since inception is negatively associated with fund flows, and the observed Average Annual Return of the self-designated Benchmark since inception is positively associated with flows in the month after being viewed. Out-performance of the fund relative to the self-designated benchmark, measured over the entire history of the fund, up to the filing date, is associated with net outflows. While aggregate results from the previous section indicate that the average EDGAR viewer is a net buyer of mutual funds, these results suggest that there exist within the EDGAR filing viewer set, investors who view prospectus filings of funds they already own, presumably to make hold or sell decisions. A large out-performance of the fund relative to the self-designated benchmark, since the fund's inception may indicate to the fund holder viewer that the fund has achieved its attainable returns, e.g., "Alpha" has been captured, especially if that out-performance, which is conveniently annualized, is larger than the out-performance of the fund within the past year. This deceleration of returns could be viewed as a negative signal, where fund holders divest and speculators pass on investment.

The component of expenses as Management Fees is most salient for investors, which is negatively related to post information acquisition fund flows. The net expenses, which is the true 'total cost' of the fund, after fee waivers and reimbursable expenses are included, is not associated with net fund flows, nor are 12b-1 fees.

5.6) Does investor response to acquired information vary with information timeliness?

Costly information acquisition is a phenomenon that keeps markets in disequilibrium in a Grossman and Stiglitz (1980) sense. Slow information diffusion is an oft cited explanation rationalizing movement of prices (e.g. post-earnings announcement drift). In canonical flow-performance models, mutual fund flows are predicted by prior period returns, as if investors have perfect information sets. If investors are acquiring, and reacting to information contained within EDGAR filings, there is an observable gap between when that information was acquired and when that information was released and current. Does the relative salience of information acquired by mutual fund investors depend on how old the information is? To put it another way, do investors respond to the acquisition of stale information? If so, at what point does information become "expired"?

To examine these questions, I'll use the previous setting of information content of EDGAR views of XBRL filings in different subsets of Filing Age, defined as the difference in days from when the filing was released to the public and when it was ultimately viewed by the investor. I use the most salient information associated with the filing from the previous analysis as independent variables: Management Fees, Average Annual Return of Fund/Benchmark in the

Past Year/Since Inception. To test if there is a clientele effect, I interact each of these independent variables with a dummy variable of if the IP address is coming from an identifiable financial organization. The dependent variable is fund flows in the month subsequent to the view of the fund(s) associated with the viewed filing. I find evidence, in Table 9, that retail (e.g. non-financial) investors are responding to acquisition of stale information through fund flows.

Among the different sub-samples based on timing of when the filing is viewed, I observe the same pattern as before: returns of fund in the past year are positively associated with post viewing fund flows and benchmark returns in the past year are negatively associated with post viewing fund flows. The opposite association is made when fund/benchmark returns are measured since inception. Returns of the fund in excess of the self-designated benchmark positively predict flows until the sample is restricted to filings older than two years. Management fees salience as a predictor of post-viewing flows, while insignificant in the recently viewed filing sample increases as the sample restricts to older files, in both statistical significance and magnitude.

Information contained within filings that are in excess of a year old is significantly correlated with fund flows in the month after the information was acquired. This is suggestive of non-Markovian behavior of mutual fund investors' response to information acquisition, supporting theories of slow information diffusion through an investor attention channel.

6) ABNORMAL ATTENTION AND RETURNS

Does Abnormal View Share predict returns? Are the attention driven investments made generating returns? In Table 10, I sort funds into portfolios rebalanced monthly based on Abnormal View Share measures and report the time series averages over the sample period from Jan 2007-June 2017. While not perfectly monotonic, the portfolios of funds with higher levels of Abnormal View Share have directional higher returns. The decile spread portfolio based on Abnormal View Share generates a 11.77 basis points return per month. Next, I perform CAPM regressions for each portfolio and the "high - low" spread portfolio and report the alpha estimates in Panel B of Table 10. The monthly decile spread portfolio sorted by abnormal view share estimates a CAPM alpha of 17.44 basis points per month. The majority of the return predictability comes from the short leg of the portfolio. A sudden absence in investor attention is a negative signal for future fund under-performance. As mutual funds are subject to short sale constraints, this result cannot be arbitraged away because funds cannot be sold short.

Next, I examine portfolios sorted on two measures of Abnormal View Share where the views are partitioned between filings viewed within 30 days of release, which I call "Recent", and those in excess of 30 days which I call "Old". Examining portfolio sorts of Recent Abnormal View Share produce more amplified results. The decile spread portfolio earns 17.08 basis points per month, and estimates a CAPM Alpha of 21.25 basis points per month, again with the majority of the source of the predictability coming from the short portfolio. Figure 2 plots the time series of the Abnormal View Share - Recent sorted spread portfolio returns. In contrast, the "Old" Abnormal View Share sorted portfolios have much more muted return results, the CAPM Alpha is of the decile spread portfolio is estimated to be 10 basis points per month.

I then further partition the sample into "Active Equity" funds, by removing CIKs that contain a majority of their assets in either bonds, index funds or ETFs. The remain funds I call "Passive", to test if attention to new filings has a strong return effects for actively managed equity funds. I find that there is indeed a larger return difference in the spread portfolio for the Active Equity sample versus the Passive sample, 21.18 to 9.08 basis points per month respectively. The Active Equity spread portfolio estimates a monthly CAPM alpha of 24 basis points per month, while the Passive spread portfolio's CAPM alpha estimate is not statistically significantly different from zero.

For Robustness, I sort portfolios based on Abnormal View Share, rebalancing once per quarter. In Table 11, the results remain qualitatively similar, with spread portfolio returns and estimated CAPM alphas roughly three times in magnitude from the monthly estimates. I also observe that the predictability in returns is also coming from the short portfolio in the quarterly case as well.

In Tables 12 and 13, I test if next periods returns can be predicted by Abnormal View Share in a panel and Fama-MacBeth settings respectively. Abnormal View Share is positively related to value weighted CIK level returns in the fixed effects panel regressions, but does not survive the full mutual fund characteristic controls specifications in the Fama MacBeth setting.

Abnormal View Share is relatively uncorrelated with other known predictors of mutual fund returns. The correlations with Active Weight (Doshi, Elkamhi, and Simutin (2015)), R-Squared (Amihud and Goyenko (2013)), and Return gap (Kacperczyk Sialm, and Zheng (2008)), are -.93%, 1.36% and 1.04% respectively.

In Table 14, I examine the return predictability of Abnormal View Share in the same sub-sample settings as in Table 7. As was also the case in the flow regressions in Table 7, Abnormal View Share is not significantly related to next month returns in majority Index Fund CIKs. However, AVS is positively related to returns in both the Active Equity and Active Bond subsamples. When the views are partitioned between Foreign IP addresses and Domestic, the Foreign measure of Abnormal View Share is unrelated to future returns, while the Domestic is positively related to returns in Active Bonds and Active Equity. Spikes in views to recently (within 30 days) released filings predict returns in the Active Bond subsample but not the Active Equity subsample. Conversely, spikes in views of older filings are associated with future returns in the Active Equity subsample, but not the Active Bond subsample. The Abnormal View Share of financial IP Addresses is associated with future returns, particularly in the actively managed funds, despite not being associated with flows in Table 7. If I assume the financial IP addresses are more sophisticated, and they are viewing filings that perform well in the future, then this elicits the question of what sort of information are these viewers obtaining if they are not seeking to subsequently invest in these funds. Cao, Du, Yang, and Zhang (2019) find evidence that some 13-F Filers, upon viewing their competitor's 13-F filings, engage in 'copycat' trades. While I do not explicitly test for this behavior in mutual funds, it is possible, given the volume of disclosure of information within mutual fund filings, such as quarterly snapshots of fund holdings, sophisticated competitors in the asset management could be trading on information acquired via EDGAR of their competitors.

7) CONCLUSIONS

Investors use EDGAR to examine mutual fund filings in addition to equity filings. Spikes in the investor attention to mutual funds predict mutual fund flows and growth in market share in the following month. This investor attention to flows relation is strongest for younger, more opaque funds. However, investors are viewing old, stale filings and appear to be making investment decisions, measured by flows to funds, with the information they ultimately uncover. A sudden absence of investor attention is a predictive negative signal for future fund under-performance.

Table 1: Top 10 of Sample

Panel A

Rank	Largest CIK by Assets (06/2017)	TNA ($M)	Number of Share Classes
1	Vanguard Index Funds	1,338,950	52
2	iShares Trust	1,011,929	262
3	Vanguard Bond Index Funds	433,324	25
4	Vanguard Chester Funds	385,161	26
5	DFA Investment Dimensions Group Inc	349,586	98
6	Vanguard Star Funds	344,965	11
7	PIMCO Funds	330,766	399
8	JPMorgan Trust II	275,524	164
9	JPMorgan Trust I	271,621	442
10	Vanguard Institutional Index Funds	269,437	4

Panel B

Rank	Largest CIK by Avg. EDGAR View Share	Avg Edgar View Share (%)
1	iShares Trust	1.15
2	PIMCO Funds	1.08
3	Northern Lights Fund Trust	0.97
4	Goldman Sachs Trust	0.84
5	Wells Fargo Funds Trust	0.66
6	VanEck Vectors ETF Trust	0.63
7	JPMorgan Trust I	0.59
8	Blackrock Funds	0.58
9	WisdomTree Trust	0.57
10	Fairholme Funds	0.54

Panel C

Rank	Largest CIK by Avg. Abnormal View Share	Avg AVS (bp)
1	Reserve Funds, NY	30.88
2	Legg Mason Partners Investment Funds, Inc.	7.17
3	Helios Select Fund Inc.	4.70
4	Hallmark Investment Series Trust	4.59
5	VanEck Vectors ETF Trust	4.59
6	Janus Investment Fund	3.90
7	Northern Lights Fund Trust IV	3.40
8	Enterprise Group of Funds Inc.	3.40
9	ETF Managers Trust	3.24
10	Russell Investment Co	3.16

Panel D

Rank	Document - Category	Views	% of Sample
1	485BPOS - XBRL, Prospectus	3,305,801	23.27
2	497 - XBRL, Prospectus	2,266,513	15.96
3	N-CSR - Reports	1,220,203	8.59
4	485APOS - Prospectus	1,212,094	8.53
5	N-Q - Reports	1,087,566	7.66
6	497K - Prospectus	934,888	6.58
7	N-CSRS - Reports	762,173	5.37
8	N-PX - Proxy Statement	489,641	3.45
9	NSAR-B - Reports	334,639	2.36
10	NSAR-A - Reports	257,605	1.81

Panel A displays the largest CIKs by Assets at the end of the sample 06/2017. Number of Share Classes correspond to the number of CRSP_Fundnos mapped to the CIK. Panel B displays the top ten CIK's in terms of average monthly EDGAR View Share over the sample from 01/2007-06/2017. Panel C displays the top ten CIK's in terms of average monthly Abnormal View Share over the sample from 01/2007-06/2017. Panel D displays the most viewed filing types of Mutual Fund Filings (filings associated with CIKs of Mutual Funds) from sample.

Table 2: CIK Summary Statistics

Variable	N	Mean	SD	p25	p50	p75
Net CIK Flow	187,736	0.0018	0.0526	-0.0129	-0.0027	0.0098
Abnormal View Share	187,430	0.0000	0.0005	-0.0001	0.0000	0.0001
Past Month Return	187,291	1.0041	0.0393	0.9926	1.0034	1.0203
Past 12 Months Return	187,342	1.0542	0.1572	1.0000	1.0466	1.1309
log TNA	187,931	7.0257	2.4010	5.5330	7.1573	8.6909
Daily Return Volatility	187,683	0.0075	0.0078	0.0021	0.0059	0.0099
Morningstar Rating	188,432	2.4497	1.6125	1	3	3.7665
No Morningstar Rating Dummy	188,432	0.2446	0.4299	0	0	0
Increase in Share Class Dummy	188,432	0.0503	0.2185	0	0	0
Decrease in Share Class Dummy	188,432	0.0290	0.1677	0	0	0
Percentage of Assets in ETFs	186,415	0.0300	0.1639	0	0	0
Percentage of Assets in Index Funds	186,201	0.0414	0.1731	0	0	0
Expense Ratio	171,695	0.0100	0.0124	0.0065	0.0096	0.0125
Amount Expense Ratio Decreased	171,274	-0.0000	0.0003	0	0	0
Amount Expense Ratio Increased	188,432	0.0000	0.0007	0	0	0
log Fund Age (months)	187,931	4.8939	0.9932	4.4998	5.1487	5.5480

Table 2 displays the value weighted fund characteristics of Mutual Fund CIKs. Flow % is the monthly net flow of assets into the CIK. Abnormal View Share is defined in section 3. Returns are raw value weighted returns of the funds within the CIK. log TNA is the natural log of the total net assets of the CIK. Daily Return Volatility is the value weighted average of the daily return standard deviation within a month for each fund within a CIK. Morningstar Rating is the value weighted measure of Morningstar Star ratings of funds within a CIK, if there are no funds rated Morningstar Rating equals zero. No Morningstar Rating Dummy equals one if the CIK has no funds rated by Morningstar. Increase in Share Class Dummy equals one if the number of share classes of the CIK increased in a given month. Decrease in Share Class Dummy equals one if the number of share classes of the CIK decreased in a given month. Percentage of Assets in ETFs is the value weighted percentage of funds within the CIK that are ETFs (1 = all etfs within CIK). Percentage of Assets in Index Fund is the value weighted percentage of funds within the CIK that are index funds (1 = all index funds within CIK). Expense Ratio is the value weighted expense ratio for the funds within the CIK. Amount Expense ratio decreased is the absolute amount expense ratios fell in the given month (all values are ≤ 0). Amount Expense Ratio increased is the amount expense ratios increased in the given month (all values are ≥ 0). log Fund Age is the natural log of the value weighted average of all fund ages in months within the CIK.

Table 3: Determinants of Abnormal View Share

Dependent Variable: Abnormal View Share (bp)	Coef.	t-stat	Economic Magnitude
Past Month Return	1.261	2.77	14.74%
Past 12 Months Return	0.179	1.04	8.43%
log TNA	0.038	5.57	26.16%
Daily Return Volatility	-1.440	-0.73	-3.29%
Morningstar Rating	0.016	1.25	7.23%
No Morningstar Rating Dummy	0.061	1.22	6.87%
Increase in Share Class Dummy	0.544	5.51	33.28%
Decrease in Share Class Dummy	0.204	1.63	10.08%
Percentage of Assets in ETFs	0.315	2.10	15.08%
Percentage of Assets in Index Funds	-0.051	-0.71	-2.56%
Expense Ratio	1.369	1.53	4.10%
Amount Expense Ratio Decreased	-54.653	-2.05	-4.59%
Amount Expense Ratio Increased	-6.246	-1.32	-1.11%
log Fund Age (months)	-0.065	-3.54	-14.98%
Constant	-1.153	-2.11	0.00%
Number of Observations	167,876		
Adj- R-Squared	0.13%		
Month FE & CIK Clustered SE	Y		

Table 3 displays monthly predictive OLS regression of Abnormal EDGAR view share as the dependent variable. Economic Magnitude is the effect on the dependent variable of a one standard deviation increase in the independent variable, ceteris paribus, scaled by the mean Abnormal View Share for a CIK.

Table 4: Explaining Mutual Fund CIK-Level Net Flows

Dependent Variable: Net CIK Flows (%)	(1)	(2)	(3)
Abnormal View Share	0.6955	0.4846	0.4927
	(2.94)	(2.06)	(2.33)
Past Month Return	0.0953	0.0955	0.0906
	(13.60)	(13.89)	(14.07)
Past 12 Months Return	0.0318	0.0272	0.0223
	(13.97)	(12.50)	(12.06)
log TNA	0.0007	-0.0004	-0.0004
	(5.35)	(-2.56)	(-3.58)
Daily Return Volatility	-0.1760	-0.0831	-0.0475
	(-4.26)	(-1.85)	(-1.20)
Morningstar Rating		0.0075	0.0065
		(23.65)	(22.52)
No Morningstar Rating Dummy		0.0283	0.0245
		(21.03)	(20.41)
Increase in Share Class Dummy		0.0040	0.0001
		(4.37)	(0.13)
Decrease in Share Class Dummy		0.0173	0.0227
		(11.30)	(14.50)
Percentage of Assets in ETFs		0.0081	0.0073
		(3.81)	(3.91)
Percentage of Assets in Index Funds		0.0006	0.0005
		(0.40)	(0.40)
Expense Ratio	0.0220	0.0452	0.0285
	(0.37)	(0.74)	(0.53)
Amount Expense Ratio Decreased		-0.8474	-0.7965
		(-1.81)	(-1.66)
Amount Expense Ratio Increased		0.2076	0.2469
		(1.09)	(1.34)
log Fund Age (months)	0.0121	-0.0101	-0.0084
	(-26.61)	(-22.5)	(-21.19)
Past Month CIK Flow (%)			0.1461
			(10.00)
Constant	-0.0715	-0.0964	-0.0913
	(-9.67)	(-12.85)	(-13.31)
Observations	171,616	169,690	169,690
Adj R-Squared	5.08%	6.90%	8.81%
Month FE & CIK Clustered SE	Y	Y	Y

Table 4 displays one month ahead predictive monthly fixed effects panel regressions for dependent variable monthly net flows of assets into a CIK scaled by assets.

Table 5: Explaining Components of Mutual Fund Asset Growth

Dependent Variable:	Change in Market Share	New Inflows (%) (NSAR)	Outflows (%) (NSAR)
Abnormal View Share	14.7061	4.2647	2.5758
	(1.79)	(1.89)	(1.32)
Past Month Return	0.9261	0.0533	-0.0669
	(7.66)	(1.93)	(-2.41)
Past 12 Months Return	0.0879	-0.0229	-0.0553
	(3.44)	(-1.38)	(-3.31)
log TNA	0.0140	0.0078	0.0088
	(3.39)	(3.15)	(3.78)
Daily Return Volatility	-3.4971	-2.5647	-2.4326
	(-5.12)	(-3.34)	(-3.44)
Morningstar Rating	0.0129	-0.0121	-0.0200
	(4.24)	(-3.78)	(-6.84)
No Morningstar Rating Dummy	0.0176	0.1671	0.1313
	(1.30)	(10.15)	(8.47)
Increase in Share Class Dummy	-0.0019	0.0337	0.0260
	(-0.09)	(3.91)	(3.08)
Decrease in Share Class Dummy	0.1999	0.0596	0.0624
	(4.84)	(5.71)	(5.83)
Percentage of Assets in ETFs	0.1306	-0.0861	-0.0747
	(1.56)	(-3.92)	(-4.59)
Percentage of Assets in Index Funds	0.2717	0.0040	0.0004
	(2.40)	(0.16)	(0.02)
Expense Ratio	0.2940	-1.6971	-1.5952
	(1.35)	(-2.3)	(-2.31)
Amount Expense Ratio Decreased	2.1387	0.9436	0.9922
	(0.88)	(0.57)	(0.66)
Amount Expense Ratio Increased	1.1696	3.1167	2.9508
	(1.02)	(1.91)	(2.07)
log Fund Age (months)	-0.0210	-0.02	-0.0059
	(-3.61)	(-3.43)	(-1.08)
Constant	-1.0231	0.1445	0.2426
	(-7.92)	(2.45)	(4.17)
Observations	169,691	159,647	159,168
Adj R-Squared	1.44%	16.39%	17.34%)
Month FE & CIK Clustered SE	Y	Y	Y

Table 5 displays one month predictive regressions for dependent variables monthly changes in CIK share of mutual fund market, new inflows to the CIK as aggregated from funds' NSAR filings as a percentage of assets, and outflows from the CIK as a percentage of assets.

Table 6: Abnormal Attention Flows - Fund Opacity

Dependent Variable: Net CIK Flow (%)	(1)	(2)	(3)
Abnormal View Share	1.3045	0.0406	0.1158
	(2.41)	(0.19)	(0.31)
Post MS Rated Dummy	-0.0020		-0.0013
	(-3.33)		(-2.30)
AVS x Post MS Rated Dummy	-1.2912		-0.1135
	(-2.15)		(-0.25)
Young Fund Dummy		0.0023	0.0052
		(2.20)	(3.24)
AVS x Young Fund Dummy		2.0715	4.8797
		(2.96)	(3.04)
Young Fund Dummy x Post MS Rated			-0.0039
			(-2.41)
AVS x Young x Post MS Rated			-4.7449
			(-2.64)
Constant	-0.0942	-0.1019	-0.1013
	(-12.45)	(-12.96)	(-12.82)
Observations	169,690	169,690	169,690
Adj R-Squared	6.93%	6.92%	6.97%
Month FE & CIK Clustered SE	Y	Y	Y
CIK Characteristic Controls	Y	Y	Y

Table 6 displays one month ahead predictive monthly fixed effects panel regressions for dependent variable monthly net flows of assets into a CIK scaled by assets. CIK Characteristic Controls consist of Past Month Return, Past 12 Months Return, log TNA, Daily Return Volatility, Morningstar Rating, No Morningstar Rating Dummy, Increase in Share Class Dummy, Decrease in Share Class Dummy, Percentage of Assets in ETFs, Percentage of Assets in Index Funds, Expense Ratio, Amount Expense Ratio Decreased, Amount Expense Ratio Increase and log Fund Age (months). Post MS Rated is a dummy variable that equals one on all CIK-months equal to and after the first fund within the CIK that did not have a Morningstar Rating as of the beginning of the sample period (January 2007) obtained a rating. Young Fund Dummy equals 1 if the value weighted fund age of the CIK is under 90 months.

Table 7: Abnormal Attention & Flows Subsample Analysis

Dependent Variable: CIK Flow	All	Active Equity	Active Bond	ETF	Index
AVS - All	0.4846	0.4798	-0.2944	1.6871	-0.0454
	(2.06)	(1.88)	(-0.63)	(1.79)	(-0.05)
AVS - Recent	0.7763	0.7763	-0.2173	0.9700	0.2282
	(1.79)	(0.70)	(-0.16)	(0.92)	(0.18)
AVS - Old	0.3756	0.5584	-0.4212	3.2848	0.2929
	(1.47)	(2.00)	(-0.95)	(1.93)	(0.31)
AVS - Domestic	0.6003	0.5391	-0.2171	2.3515	0.262
	(2.02)	(1.68)	(-0.39)	(2.03)	(0.26)
AVS - Foreign	0.1656	4.101	-0.7644	1.4129	-2.0191
	(0.47)	(1.16)	(-0.86)	(0.63)	(-1.22)
AVS - Retail	0.5192	0.529	-0.3192	1.8388	-0.0817
	(2.09)	(1.95)	(-0.65)	(1.85)	(-0.09)
AVS - Financial	-0.1679	0.0022	-0.3362	0.5465	0.7445
	(-0.18)	(0.00)	(-0.16)	(0.11)	(0.19)
Observations	169,690	89,416	71,143	4,938	5,796
CIK Controls	Y	Y	Y	Y	Y
Monthly FE & CIK Clusted SE	Y	Y	Y	Y	Y

Table 7 displays one month ahead predictive monthly fixed effects panel regressions for dependent variable monthly net flows of assets into a CIK scaled by assets. CIK Characteristic Controls consist of Past Month Return, Past 12 Months Return, log TNA, Daily Return Volatility, Morningstar Rating, No Morningstar Rating Dummy, Increase in Share Class Dummy, Decrease in Share Class Dummy, Percentage of Assets in ETFs, Percentage of Assets in Index Funds, Expense Ratio, Amount Expense Ratio Decreased, Amount Expense Ratio Increase and log Fund Age (months). Each coefficient corresponds to a separate regression. CIK Characteristic Controls consist of Past Month Return, Past 12 Months Return, log TNA, Daily Return Volatility, Morningstar Rating, No Morningstar Rating Dummy, Increase in Share Class Dummy, Decrease in Share Class Dummy, Percentage of Assets in ETFs, Percentage of Assets in Index Funds, Expense Ratio, Amount Expense Ratio Decreased, Amount Expense Ratio Increase and log Fund Age (months). AVS - All is Abnormal View Share as defined in section 3. AVS - Recent is Abnormal View Share where only views to filings that are viewed within 30 days of release are counted. AVS - Old is Abnormal View Share where only views to filings that are viewed in excess of 30 days of release are counted. AVS - Domestic is Abnormal View Share where only views from IP addresses within the United States are counted. AVS - Foreign is Abnormal View Share where only views from IP addresses outside of the United States are counted. AVS - Financial is Abnormal View Share where only views from IP addresses that are from identifiable financial organizations are counted. AVS - Retail is Abnormal View Share where only views from IP addresses that are not from identifiable financial organizations are counted. Active Equity subsample is constructed by removing CIK observations that contain a majority of assets in bonds, index funds, or ETFs. Active Bond subsample is constructed by removing CIK observations that contain a majority of assets in index funds, or ETFs and having majority of assets in bonds. ETF subsample consists of CIKs that have the majority of their assets in ETFs. Index subsample consists of CIKs that have the majority of their assets in Index Funds.

Table 8: Filing Information Content and Investor Flow Response

Dependent Variable: Net Fund Flows (%) in Month After Filing View	(1)	(2)	(3)
Avg Annual Return of Fund Past Year on Filing		0.0597	0.0746
		(2.13)	(2.18)
Avg Annual Return of Fund Past 5 Years on Filing		0.0284	0.0717
		(0.33)	(0.75)
Avg Annual Return of Fund Past 10 Years on Filing		-0.1934	-0.2334
		(-3.76)	(-3.28)
Avg Annual Return of Fund Since Inception on Filing		-0.1139	-0.1102
		(-2.19)	(-2.13)
Avg Annual Return of Benchmark Past Year on Filing		0.0819	0.0988
		(-2.51)	(-2.35)
Avg Annual Return of Benchmark Past 5 Years on Filing		0.0263	-0.017
		(0.39)	(-0.22)
Avg Annual Return of Benchmark Past 10 Years on Filing		0.1228	0.137
		(2.88)	(2.42)
Avg Annual Return of Benchmark Since Inception on Filing		0.0932	-0.0813
		(2.14)	(1.73)
Management Fees on Filing (%)	-1.1533		-3.9489
	(-2.63)		(-3.46)
12b-1 Fees on Filing (%)	-0.7139		-1.1974
	(-1.39)		(-1.16)
Net Expenses on Filing (%)	0.2124		0.3948
	(2.3)		(1.05)
Past Month Return	-0.1327	-0.1393	-0.1430
	(-4.97)	(-3.17)	(-2.75)
Past 12 Months Return	0.0713	0.0535	0.0656
	(8.25)	(5.08)	(5.07)
Morningstar Rating	0.0096	0.0084	0.0103
	(10.42)	(3.98)	(4.06)
log Total Net Assets	-0.0014	0.0005	0.0013
	(-2.35)	(0.59)	(1.04)
Past Month Daily Volatility	-0.0769	0.1046	0.5876
	(0.43)	(0.28)	(1.40)
log Fund Age	-0.0128	-0.0056	-0.0102
	(-11.38)	(-3.76)	(-3.17)
Expense Ratio as of Past Month	0.0135	0.4359	2.6988
	(0.04)	(0.97)	(2.63)
Constant	0.1096	0.0987	0.0881
	(4.6)	(2.33)	(1.77)
Observations	216,291	109,356	81,155
Adj R-Squared	5.30%	6.23%	7.70%
Month FE & CIK Cluster SE	Y	Y	Y

Table 8 displays one month ahead predictive fixed effects panel regression for dependent variable monthly net flows of assets into funds associated with a filing scaled by assets. Past Month Return, Past 12 Months Return, Morningstar Rating, log Total Net Assets, Past Month Daily Volatility, log Fund Age, and Expense Ratio as of Past Month are value weighted measures of all funds associated with a viewed filing of forms 497 or 485BPOS from January 2011-June 2017, on the month it was viewed. All other independent variables are obtained from the Mutual Fund Prospectus Risk/Return Summary Data Set[21] from the associated filing that is viewed by an EDGAR user. The unit of observation is a form 497 or form 485BPOS view during January 2011-June 2017.

[21] https://www.sec.gov/dera/data/mutual-fund-prospectus-risk-return-summary-data-sets

Table 9: Filing Age and Salient Information

Dep Var: Net Fund Flows (%) in Month After View Age of filing when viewed	All	Recent Views	Exceeds 30 Days	Exceeds 180 Days	Exceeds 365 Days
Avg Ret of Fund Past Year on Filing	0.0288	0.0899	0.0205	0.0128	0.0076
	(1.73)	(2.37)	(1.55)	(1.12)	(0.7)
Avg Ret of Fund Since Inception on Filing	-0.0622	-0.1156	-0.0540	-0.0563	-0.0570
	(-2.19)	(-2.00)	(-2.13)	(-2.24)	(-2.54)
Avg Ret of Benchmark Past Year on Filing	0.0357	-0.1402	-0.0209	-0.0102	-0.0020
	(-1.91)	(-2.74)	(-1.52)	(-0.83)	(-0.16)
Avg Ret of Benchmark Since Inception on Filing	0.0552	0.1181	0.0431	0.0387	0.0318
	(2.11)	(2.34)	(1.89)	(1.81)	(1.59)
Management Fees on Filing (%)	-1.2851	-0.5715	-1.3190	-1.3898	-1.5328
	(-2.82)	(-0.78)	(-3.04)	(-3.61)	(-4.10)
Financial View	-0.0039	-0.0084	-0.0024	0.0023	-0.0024
	(-1.44)	(-1.22)	(-1.03)	(-1.00)	(-0.91)
Financial x Avg Ret of Fund Past Year on Filing	-0.0117	-0.0385	-0.0108	0.0121	-0.0392
	(-0.6)	(-0.79)	(-0.61)	(-0.56)	(-1.21)
Financial x Avg Ret of Fund Since Inception on Filing	0.0391	0.1271	-0.0125	0.0070	0.0513
	(1.11)	(1.27)	(0.44)	(-0.18)	(1.38)
Financial x Avg Ret of Benchmark Past Year on Filing	0.0135	0.0393	-0.0161	0.0158	0.0419
	(0.63)	(0.71)	(0.84)	(0.67)	(1.19)
Financial x Avg Ret of Benchmark Since Inception on Filing	-0.0230	-0.0347	-0.0171	-0.0009	-0.0319
	(-0.68)	(-0.39)	(-0.56)	(-0.02)	(-0.97)
Financial x Management Fees on Filing (%)	0.1648	0.2981	0.1268	0.1296	-0.2035
	(0.54)	(0.40)	(0.45)	(0.43)	(-0.54)
Past Month Ret	-0.1374	-0.2637	-0.1108	-0.1063	-0.1243
	(-4.47)	(-2.79)	(-4.10)	(-3.89)	(-3.96)
Past 12 Months Ret	0.0590	0.0860	0.0550	0.0519	0.0436
	(6.16)	(3.79)	(5.75)	(4.72)	(3.59)
Morningstar Rating	0.0095	0.0102	0.0091	0.0087	0.0088
	(8.89)	(4.45)	(9.88)	(9.81)	(9.07)
log Total Net Assets as of Past Month	-0.0013	-0.0034	-0.0008	-0.0007	-0.0004
	(-2.16)	(-2.86)	(-1.65)	(-1.68)	(-0.85)
Past Month Daily Volatility	0.1654	0.6807	0.0962	0.1631	0.0874
	(0.74)	(1.20)	(0.52)	(0.81)	(0.36)
log Fund Age	-0.0129	-0.0145	-0.0113	0.0101	-0.0101
	(-9.06)	(-4.24)	(-9.60)	(-8.32)	(-7.11)
Expense Ratio as of Past Month	0.1573	-0.7984	0.3461	0.3587	0.4547
	(0.45)	(-1.29)	(1.10)	(1.26)	(1.56)
Constant	0.1278	0.2638	0.0907	0.0821	0.1063
	(4.49)	(2.97)	(3.78)	(3.63)	(3.74)
Observations	173,240	32,825	140,295	102,543	67,896
Adj R-Squared	6.09%	5.54%	6.83%	7.70%	7.75%
Month FE & CIK Cluster SE	Y	Y	Y	Y	Y

Table 9 displays one month ahead predictive fixed effects panel regression for dependent variable monthly net flows of assets into funds associated with a filing scaled by assets. Past Month Return, Past 12 Months Return, Morningstar Rating, log Total Net Assets, Past Month Daily Volatility, log Fund Age, and Expense Ratio as of Past Month are value weighted measures of all funds associated with a viewed filing of forms 497 or 485BPOS from January 2011-June 2017, on the month it was viewed. All other independent variables are obtained from the Mutual Fund Prospectus Risk/Return Summary Data Set from the associated filing that is viewed by an EDGAR user. Columns subset the sample to include filings viewed within 30 days for "Recent Filings", or filings that are viewed after a specified number of days since the release of the filing. The unit of observation is a form 497 or form 485BPOS view during January 2011-June 2017.

Table 10: Monthly Decile Portfolios sorted on Abnormal View Share

Panel A Avg Returns	(bp)	AVS	- All	AVS -	Recent	AVS	- Old	AVS - Active	Recent Equity	AVS - Passive	Recent
low AVS	1	35.00		32.49		37.55		29.43		37.59	
	2	41.43		38.27		40.53		44.93		29.14	
	3	44.79		34.88		42.34		29.20		46.03	
	4	44.83		40.99		47.44		36.78		45.00	
	5	45.95		45.04		41.52		43.82		57.31	
	6	41.96		44.18		48.06		60.74		30.11	
	7	44.75		46.64		50.30		40.74		53.60	
	8	43.88		54.20		46.04		52.92		49.22	
	9	52.14		48.55		50.64		44.83		51.29	
high AVS	10	46.77		49.57		42.63		50.61		46.67	
	H-L	11.77		17.08		5.08		21.18		9.08	

Panel B CAPM Alpha	(bp)	AVS Est	- All t-stat	AVS - Est	Recent t-stat	AVS Est	- Old t-stat	AVS - Active Est	Recent Equity t-stat	AVS - Passive Est	Recent t-stat
low AVS	1	-12.78	-1.90	-14.37	-1.98	-9.42	-1.49	-17.82	-2.46	-8.16	-0.81
	2	-5.72	-0.83	-7.54	-0.96	-6.48	-1.04	-3.86	-0.51	-3.85	-0.36
	3	1.12	0.20	-7.65	-1.10	-1.00	-0.13	-17.24	-2.48	13.65	1.22
	4	2.53	0.33	0.25	0.04	6.64	0.93	-9.66	-1.18	18.64	1.73
	5	8.77	1.22	5.08	0.68	3.36	0.48	2.03	0.25	31.15	2.54
	6	3.61	0.59	0.72	0.09	8.39	1.26	15.82	1.80	-1.18	-0.10
	7	0.89	0.15	5.84	0.95	7.38	0.94	3.39	-0.45	22.71	2.16
	8	2.04	0.30	11.85	1.77	2.07	0.31	6.62	0.88	17.42	1.53
	9	6.66	1.06	5.15	0.75	6.40	0.98	-0.53	-0.07	19.35	1.86
high AVS	10	4.66	0.66	6.89	1.06	0.58	0.09	6.21	0.93	6.36	0.69
	H-L	17.44	2.84	21.25	2.88	10	1.79	24.03	2.98	14.52	1.22

Table 10 Panel A sorts CIKs into monthly decile portfolios based on Abnormal View Share and reports average returns in basis points. Panel B reports Alphas and associated t-statistics from a CAPM OLS Regression. AVS - All are portfolios sorted on Abnormal View Share as defined in section 3. AVS - Recent are portfolios sorted on Abnormal View Share where only views to filings that are viewed within 30 days of release are counted. AVS - Old are portfolios sorted on Abnormal View Share where only views to filings that are viewed in excess of 30 days of release are counted. Active Equity subsample is constructed by removing CIK observations that contain a majority of assets in bonds, index funds, or ETFs. Passive subsample consists of all assets that are majority of assets in bonds or majority index funds.

Table 11: Quarterly Decile Portfolios sorted on Abnormal View Share

Average Returns (bp)		AVS	All	AVS	Recent
low AVS	1	104.34		95.34	
	2	116.69		109.8	
	3	115.77		108.48	
	4	137.46		132.84	
	5	114.71		132.32	
	6	122.16		137.88	
	7	105.05		129.29	
	8	139.03		159.37	
	9	179.94		152.31	
high AVS	10	143.88		146.97	
	H-L	39.54		51.63	

CAPM Alpha (bp)		AVS Estimate	All t-stat	AVS Estimate	Recent t-stat
low AVS	1	-44.72	-2.05	-44.43	-1.98
	2	-32.85	-1.51	-30.39	-1.20
	3	-22.90	-1.20	-20.45	-0.82
	4	14.90	0.56	0.18	0.01
	5	4.85	0.20	10.99	0.41
	6	14.94	0.82	-5.98	-0.27
	7	-36.53	-1.47	7.05	0.35
	8	5.35	0.26	20.19	0.78
	9	40.66	1.55	17.42	0.82
high AVS	10	9.13	0.33	16.45	0.69
	H-L	53.85	2.29	60.88	2.45

Table 11 Panel A sorts CIKs into decile portfolios once per quarter based on Abnormal View Share and reports average returns in basis points. Panel B reports Alphas and associated t-statistics from a CAPM OLS Regression. AVS All are portfolios sorted on Abnormal View Share as defined in section 3. AVS Recent are portfolios sorted on Abnormal View Share where only views to filings that are viewed within 30 days of release are counted.

Table 12: Panel Regression

Dependent Variable: CIK Return			
Abnormal View Share	0.5594	0.4979	0.4710
	(2.79)	(2.64)	(2.36)
Past Month Return		0.1188	0.1107
		(14.11)	(13.89)
Past 12 Months Return			0.0003
			(0.30)
log TNA			-0.0001
			(-1.74)
Daily Return Volatility			-0.0799
			(-2.37)
Morningstar Rating			0.0002
			(1.75)
No Morningstar Rating Dummy			-0.0022
			(-3.84)
Increase in Share Class Dummy			-0.0003
			(-1.08)
Decrease in Share Class Dummy			-0.0003
			(-0.82)
Percentage of Assets in ETFs			0.0018
			(3.17)
Percentage of Assets in Index Funds			0.0007
			(1.70)
Expense Ratio			-0.0588
			(-1.13)
Amount Expense Ratio Decreased			0.8405
			(3.22)
Amount Expense Ratio Increased			-0.1684
			(-1.52)
log Fund Age (months)			0.0001
			(0.72)
Constant	1.0042	0.8849	0.8938
	(14345)	(104.52)	(108.69)
Observations	187,332	187,193	196,960
Adj R-Squared	52.00%	52.68%	51.83%
Month FE & CIK Clustered SE	Y	Y	Y

Table 12 displays one month ahead predictive monthly fixed effects panel regressions for dependent variable of value weighted CIK returns.

Table 13: Fama-MacBeth Regressions

Dependent Variable: CIK Returns			
Abnormal View Share	0.4081	0.3199	0.0341
	(1.99)	(1.95)	(0.28)
Past Month Return		0.0645	0.0525
		(1.44)	(1.81)
Past 12 Months Return			0.0179
			(2.39)
log TNA			-0.0000
			(-0.52)
Daily Return Volatility			0.0796
			(0.34)
Morningstar Rating			0.0000
			(0.09)
No Morningstar Rating Dummy			-0.0012
			(2.22)
Increase in Share Class Dummy			0.0001
			(0.26)
Decrease in Share Class Dummy			-0.0001
			(-0.27)
Percentage of Assets in ETFs			0.0007
			(0.98)
Percentage of Assets in Index Funds			-0.0001
			(-0.28)
Expense Ratio			-0.0781
			(-2.96)
Amount Expense Ratio Decreased			0.1730
			(0.45)
Amount Expense Ratio Increased			-0.1069
			(-0.15)
log Fund Age (months)			0.0001
			(1.02)
Constant	1.0042	0.9390	0.9315
	(401.05)	(20.69)	(33.2)
Observations	187,332	187,193	196,960
Adj R-Squared	0.11%	18.7%	47.37%

Table 13 displace the results of the Fama-MacBeth two-step procedure with t-statistics in parentheses calculated from Newey-West adjusted 2 lag standard errors.

Table 14: Abnormal Attention & Return Subsample Analysis

Dependent Variable: CIK Return	All	Active Equity	Active Bond	ETF	Index
AVS - All	0.4710	0.3167	0.6156	0.4334	0.0814
	(2.36)	(2.02)	(2.09)	(0.77)	(0.21)
AVS - Recent	0.9276	0.1927	1.9309	0.3622	0.2069
	(3.34)	(0.74)	(2.82)	(0.68)	(0.40)
AVS - Old	0.2566	0.3551	0.4133	0.6604	-0.0371
	(1.04)	(2.22)	(1.00)	(0.87)	(-0.06)
AVS - Domestic	0.6097	0.4046	0.7501	0.5865	0.1197
	(2.42)	(1.94)	(2.17)	(0.97)	(0.24)
AVS - Foreign	0.2893	0.2178	0.3749	0.7829	0.3244
	(1.36)	(1.20)	(0.99)	(0.49)	(0.55)
AVS - Retail	0.4434	0.3035	0.5536	1.8388	0.1398
	(2.02)	(1.79)	(1.66)	(1.85)	(0.32)
AVS - Financial	1.4148	0.8279	2.4326	0.5465	-1.1522
	(2.97)	(1.72)	(2.43)	(0.11)	(-1.06)
Observations	169,690	89,416	71,143	4,938	5,796
CIK Controls	Y	Y	Y	Y	Y

Table 14 displays one month ahead predictive monthly fixed effects panel regressions for dependent variable monthly CIK Returns. CIK Characteristic Controls consist of Past Month Return, Past 12 Months Return, log TNA, Daily Return Volatility, Morningstar Rating, No Morningstar Rating Dummy, Increase in Share Class Dummy, Decrease in Share Class Dummy, Percentage of Assets in ETFs, Percentage of Assets in Index Funds, Expense Ratio, Amount Expense Ratio Decreased, Amount Expense Ratio Increase and log Fund Age (months). Each coefficient corresponds to a separate regression. CIK Characteristic Controls consist of Past Month Return, Past 12 Months Return, log TNA, Daily Return Volatility, Morningstar Rating, No Morningstar Rating Dummy, Increase in Share Class Dummy, Decrease in Share Class Dummy, Percentage of Assets in ETFs, Percentage of Assets in Index Funds, Expense Ratio, Amount Expense Ratio Decreased, Amount Expense Ratio Increase and log Fund Age (months). AVS - All is Abnormal View Share as defined in section 3. AVS - Recent is Abnormal View Share where only views to filings that are viewed within 30 days of release are counted. AVS - Old is Abnormal View Share where only views to filings that are viewed in excess of 30 days of release are counted. AVS - Domestic is Abnormal View Share where only views from IP addresses within the United States are counted. AVS - Foreign is Abnormal View Share where only views from IP addresses outside of the United States are counted. AVS - Financial is Abnormal View Share where only views from IP addresses that are from identifiable financial organizations are counted. AVS - Retail is Abnormal View Share where only views from IP addresses that are not from identifiable financial organizations are counted. Active Equity subsample is constructed by removing CIK observations that contain a majority of assets in bonds, index funds, or ETFs. Active Bond subsample is constructed by removing CIK observations that contain a majority of assets in index funds, or ETFs and having majority of assets in bonds. ETF subsample consists of CIKs that have the majority of their assets in ETFs. Index subsample consists of CIKs that have the majority of their assets in Index Funds.

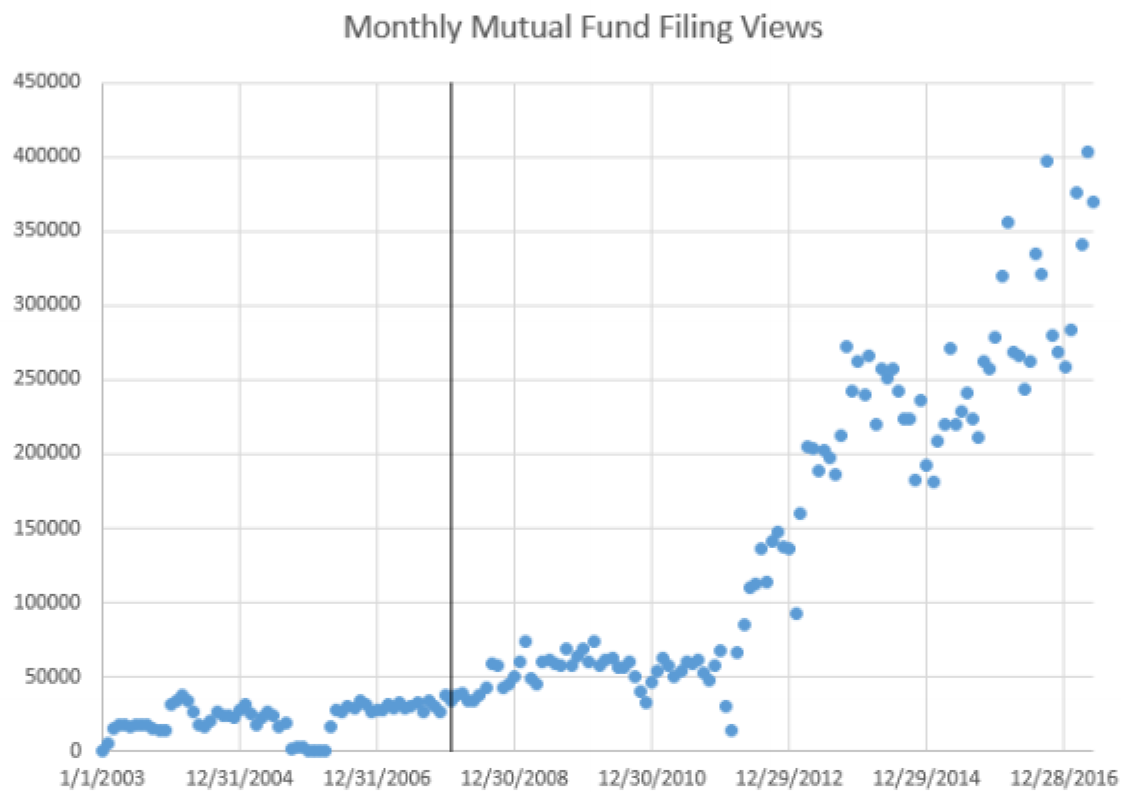

Figure 1: Restricted EDGAR Log File to include only Mutual Fund CIKs. I removed Robot Downloads according to Ryans (2017). The sample consists of 18.5M observations of Mutual Fund filing views.

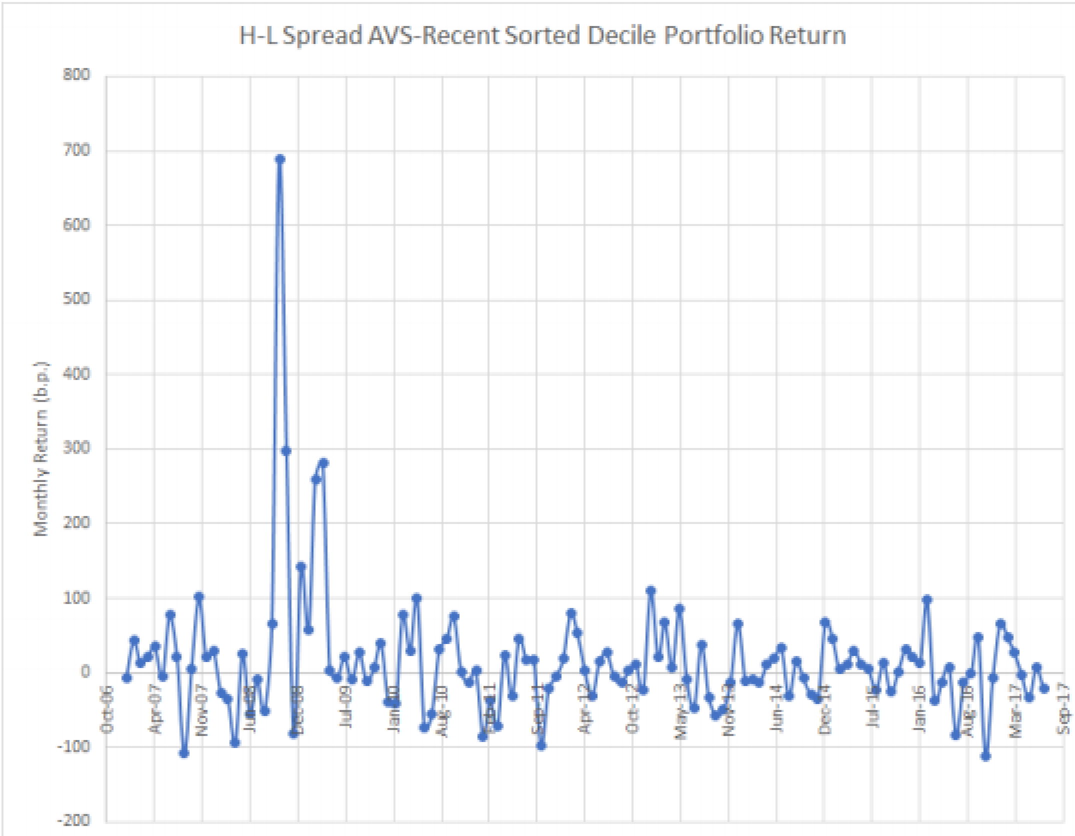

Figure 2: Monthly time series of H-L decile spread portfolio returns sorted by Abnormal View Share - Recent.

CHAPTER 2

1) INTRODUCTION

Attention is bounded and costly (Kahneman 1973) for individual investors. In contrast to other research relating proxies to the supply side of investor attention, e.g. news about fund (Sirri and Tufano 1998), or components of fund holdings (Solomon, Soltes, and Sosyura 2014) to equity mutual fund flows, I examine a proxy for the demand side of investor attention (spikes in search volume of mutual fund tickers on Google, obtained from Google Trends). If there is an event which is informative, only those paying attention will react. This idea has spurned a burgeoning literature in limited investor attention and its effects on slow information diffusion into asset prices, news reaction and consumer behavior.[22] I use a direct and unambiguous measure of investor attention utilizing publicly available data from Google Trends. Google Trends data allows me to examine a revealed preference for the demand of information by investors, examining the propensity for search queries of keywords over time. Google is the predominant way individual search for information on the Internet both in the United States[23], and worldwide[24]. Google Trends has been used as a proxy for investor attention, beginning with Da, Engelberg, and Gao (2011) (DEG 2011 hereafter)[25]. Investor attention research is related to work in News and Media, but it gains insight from the demand side of information as opposed to the supply side. Huberman and Regev (2001) famously demonstrate an example where investors only react to news if they pay attention to it. In a seminal paper, Merton (1987) develops a model of capital market equilibrium around the idea that investors do not invest in assets to which they do not pay attention. Solomon, Soltes and Sosyura (2014) relate media coverage of fund holdings and its

---

[22] See Hirshleifer and Teoh (2003), Grullon, Kanatas and Weston (2004), Tetlock (2007) (2010), Tetlock, Saar-Tsechansky and Macskassy (2008), Barber and Odean (2008), Fang and Peress (2009), Hirshleifer Lim, and Toeh (2009), Da, Engelberg, and Gao (2015), Fisher, Martineau and Sheng (2016), among many others.

[23] Google has maintained servicing the majority of search queries in the United States over time. https://www.statista.com/statistics/267161/market-share-of-search-engines-in-the-united-states/

[24] Google has been averaging around a 90% Desktop market share of search over time. https://www.statista.com/statistics/216573/worldwide-market-share-of-search-engines/

[25] Google Trends data for individual equity tickers has also been used by Preis, Reith, and Stanley (2010), Preis, Moat, and Stanley (2013), Ben-Rephael, Da, and Israelsen (2016), among others

impact on fund flows, claiming high-visibility equity winners attract more flows to the funds that hold them than winners that do not have as much media coverage.

Considerably less attention has been paid to investor attention within the mutual fund market, a gap in the literature this paper hopes to fill. Jain and Wu (2000) argue that Mutual Fund advertising in print media signals superior skill to investors, resulting in increased inflows but not superior performance. The goal of advertising is to inform, persuade or remind customers about your brand or product, thereby attracting attention. Sirri and Tufano (1998) suggest that due to costly search, the flow to performance relationship is most pronounced in funds that have highest fees, their proxy for unobserved marketing effort. They hypothesize that if consumers can collect costless information about mutual funds, they would expect to find a flow to performance relationship amongst the best/worst performing funds and negative relationships between flows with respect to fees and risk. Barber, Odean and Zheng (2005) argue that if investors were focusing on minimizing search costs they would be holding funds more than trading, but they note in their brokerage data from 1987-1993 roughly half of all mutual fund sales are followed by purchases within 3 weeks. They argue instead that in their data, mutual fund marketing does work, as measured by those funds employing 12b-1 fees. Since the dot-com boom, technology has dramatically lowered search costs of individuals for information. This paradigm shift in reduction of information search costs combined with the ability to examine revealed investor attention with Google searches[26], of which I find the funds with 12b-1 fees have lower levels, and decreasing levels over time of attention, even though the 12b-1 fees are collected with the expressed purpose of marketing the individual funds.

DEG 2011 propose that Google search activity is a better measure of investor attention than indirect proxies such as news, advertising expenditure, and volume, as it is a revealed attention measure. If an individual is searching for something on Google, they are certainly paying attention to it. DEG 2011, and the related literature utilizing Google Trends as a proxy for investor

[26] It should also be noted accessing this information is costless to the investor in terms of real dollars when searching on Google.

attention of specific assets seek to identify changes or spikes, "Abnormal Investor Attention", within the time series of Google search volume of an individual query. While this is an informative measure of changes in attention, previous researchers are unable to distinguish between the relative amount of search volume from two different tickers, and thus can only make inferences regarding a "standard deviation" increase in the search volume for one specific ticker relative to searches for the same ticker at different points in time. In addition to examining Google Trends data in the mutual fund market, I develop a new methodology which allows for measurement of "Absolute Attention" levels, as opposed to changes within the cross-section of funds. This new measure should work hand in hand with measures of changes to find new insights into future investor attention research utilizing Google search volumes.

My alternative measure for 'Absolute Attention' leverages the ability for Google Trends to measure multiple (up to 5) queries simultaneously, and be able to make cross-sectional comparisons between different queries. I can distinguish if two different mutual funds have search volumes on Google at the same time which are several orders of magnitude apart. However this measure of level of attention, a quantity that is proportional to actual number of queries between mutual funds comes at the cost of being unable to capture as fully the variation in attention in the time series. The relative scaling of multiple queries employed to compare many different queries relative to the same absolute benchmark causes a coarsening of the variation over time series of the reported search volumes with respect to an individual query. To illustrate the differences in approaches, consider Figures 3 and 4, which is how DEG 2011 and previous papers have examined Google trends data, capturing the full possible variation in the time series, for two different mutual funds, 'BLUEX' and 'PTTAX', each scaled by their maximum at 100 'Interest over time', what DEG 2011 call the Search Volume Index, or SVI. For 'PTTAX' this maximal scaling occurs at December 2010. Where this paper adds to the existing literature is in Figure 5, where it becomes possible to compare the search volumes of different strings at the same time. In August 2009, 'BLUEX' had twice as much search volume as 'PTTAX', as designated by the 100 vs. 50 in the raw SVI data. The series for 'PTTAX' retains its shape, but now at December 2010 produces a value of 69 instead of 100. Thus the range of possible variation shrinks from 1-100 to 1-69, a cost at being able to compare the two series together. Whichever series maintains maximal search content will retain its same shape as if it was searched alone, and for this reason, time series spikes in searches will be more statistically powerful when measured in the naive SVI way of DEG 2011.

By aggregating my individual 'Absolute Attention' measure of tickers, I create point estimates for relative levels investor attention between categories of funds, a statistic that has not be documented before to my knowledge. Across the universe of mutual funds, index funds[27] are 2.5 more likely to be searched for than non-index funds. Institutional funds[28], are 22.7 times more likely to be searched for than non-index funds, with half of all levels of investor attention concentrated in Domestic Equity Institutional funds. I find that fund size[29] is a strong positive predictor of more Absolute Attention, but a negative predictor of Abnormal Investor Attention. Sicherman, Loewenstein, Seppi and Utkus (2016) study financial attention of Vanguard investors, observing for this set of investors who hold equity positions pay more attention to their investments in rising stock markets than in falling markets and attention decreases in market volatility. I find daily return volatility for an individual month negatively predicts Abnormal Investor Attention but positively predicts Absolute Attention in the following month. Fees negatively predict Absolute Attention, but appear insignificant in regressions for Abnormal Investor Attention, suggesting investors have learned to avoid paying attention to funds with high fees. Adding current levels of Absolute Attention as a control for Abnormal Investor Attention more than doubled the R-squared in my largest model specification.

Moreover, by utilizing searches for *individual* mutual fund tickers, Google Search volume is a close proxy for aggregate focused attention on a specific mutual fund at a given point in time. I show that this focused attention has a magnifying impact on flows, with 69% of total inflows in the mutual fund market per month are concentrated in less than 3% funds per month, those with

[27] Indicator that equals 1 if any of the categories of findex_fund_flag in CRSP Survivor-Bias-Free Mutual Fund Database are present.
[28] Indicator that equals 1 if fhinst_fund is labeled 'Y' in CRSP
[29] Measured by log(Total Net Assets)

the highest attention and exceeding the bottom 30% of returns per month. Past returns are a strong predictor of flows, and R-squares increase by more than double when including either of the attention variables as a control. I don't find support for the hypothesis of convexity of flows relative to performance in the data, rather, the squared returns in my regressions all have a negative signs (indicating concavity in returns), with limited levels of significance. In non-parametric sorts, amongst high attention funds, there does appear to be some visual evidence of this relation, with average flows being flat amongst funds with attention levels below Google's threshold search level. In a horse-race between Absolute Attention and Abnormal Investor Attention, the latter wins in predictive regressions for dollar flows (measured as in Frazzini and Lamont (2008)), and Changes in Market Share (measured as in Spiegel and Zhang (2013)). Changes in attention appear to be more important to fund flows than level of attention.

2) DATA

In contrast to DEG 2011, and other studies that utilize Google Trends data to describe investor attention in the equity market, I focus on the investor attention in the mutual fund market. I examine individual queries for all Mutual Fund NASDAQ tickers that appear in CRSP survivorship bias-free mutual fund data from January 2004-March 2017. The data come from CRSP survivorship bias-free mutual fund data set and Google Trends. Google Trends[30] data is an unbiased sample of Google search data, and can be thought of a representative sample of the search behavior of an individual query for the general population. I attempt to collect Google Trends data for all Mutual Fund NASDAQ tickers that appear in CRSP survivorship bias-free mutual fund data from January 2004-March 2017. Google searches act as a proxy for the attention that investors[31] are revealing at a given point in time, as a search for an individual mutual fund implies one is paying attention to it. Google search volume are representative of

[30] http://www.google.com/trends

[31] This has been hypothesized as predominantly retail, non-institutional investor attention by DEG (2011) and Ben-Rephael, Da, and Israelsen (2016)

collective attention as it is the search engine which receives the majority of queries in the United States[32] and worldwide[33].

The raw Google Trends data that I have collected is different than how it appeared in 2009, e.g., DEG 2011's dataset, as Google has adapted the functionality of the product. In DEG 2011, the search volume index was only available weekly and was the number of searches for a individual query scaled by its time series average, to two decimal places. Over time, the product has evolved to capture search data for different frequencies. Presently the product now reports, under "Interest over time", for each queried keyword(s), an integer measure from 0-100, where 100 denotes the largest amount of searches that occurred within the observed frequency of the queried time frame. Each data point is divided by the total searches of the geography and time range queried to compare to relative popularity over time. A query in Google Trends from "2004-Present" (the earliest possible date) executed today will report a monthly time frame, and the raw data of 100 indicates the month which the queried keyword(s) were observed to be more searched relative to all other months. When the window is shorter the product allows higher frequency data to be collected, which is of interest for future research. Query windows of 3 months or less will report a daily measure, where 100 corresponds to the day at which the search was most frequent. I have collected monthly data for utilizing the 2004-Present time frames, and fixing the geography to be worldwide as well as daily data for the funds that appear in the monthly sample. As the Google Trends data is itself a sample, it is important to note that some of the search data is excluded: searches that were duplicates (made by the same IP address over a short period of time), searches containing special characters, and most importantly, low volume searches. The search volume is scaled linearly to from 100 to 1. A search volume score of 0 denotes a censored value, below 1% of the maximum given the query date range/frequency. DEG 2011 also run into this truncation issue, but argue this biases against finding a significant result.

---

[32] https://www.statista.com/statistics/267161/market-share-of-search-engines-in-the-united-states/
[33] https://www.statista.com/statistics/216573/worldwide-market-share-of-search-engines/

In order for a fund to be included in my sample, Google Trends must report a non-zero value for the search volume for the fund's NASDAQ ticker in at least one month. The monthly sample coverage is plotted in Figure 6. In each time period, there is roughly 25% of the funds available each month in CRSP Survivor-bias-free US Mutual Fund data represented in the Google Trends data. As these are the most searched for funds, they also tend to be the largest. Across total net assets of the mutual fund industry, the monthly sample covers roughly 70% of the assets each month. Within the sample of 10,783 ticker queries that contain data according to the previous remark, 53.7% of the months have an observation of the truncated value of zero. The daily data is even sparser. If the three-month window used to query the Google Trend data was all zero at the monthly level, Google will not report daily observations. Across the 242,963 ticker-months that appear in my Google Trends daily sample, 72.7% of all days contain a zero observation[34]. This indicates that the majority of mutual fund search levels are roughly close to the Google Trends censored boundary.

3) METHODOLOGY

For every ticker that appears in CRSP Survivor-bias-free US Mutual Fund data from January 2004 to March 2017, I first collect data at the monthly frequency of the individual query of the mutual fund NASDAQ ticker symbol. DEG 2011's main measure is Abnormal Investor Attention, corresponding to the weekly search volume in excess of the median of the previous 8 weeks. From these monthly data I construct the equivalent measure of Abnormal Investor Attention as the log search volume in excess of the log of the average of the previous two months (roughly 8 weeks) search volume for an individual fund ticker. Define SV_m^i as the search volume of ticker *i* in month *m*, so my measure of Abnormal Investor Attention is defined as:

$$AIA_m^i = \log(SV_m^i) - \log(\frac{SV_{m-1}^i + SV_{m-2}^i}{2})$$

[34] Note, these results are before any merging to CRSP, meaning that this includes data for mutual funds which may have not existed yet, or have ceased to exist included.

I also collect data from Google Trends again using the same NASDAQ ticker symbol query if the query appeared in my sample of 10,783, but in a 3 month windows beginning at the month where trading data first appears in CRSP, which produces daily observations, scaled to the specific day within the series that has the largest search volume.

In order to obtain a measure which more closely approximates actual search volume levels, I develop a novel method of querying and transforming Google Trends data to get more than two magnitudes of variation in search volumes between individual queries. This method is general to comparisons of any large amount of Google Trends queries, not limited to mutual fund ticker applications. My alternative measure for Absolute Attention leverages the ability for Google Trends to measure multiple (up to 5) queries simultaneously, and be able to make cross-sectional comparisons between different queries. I can distinguish if two different mutual funds have search volumes on Google at the same time which are several orders of magnitude apart. However, this measure of level of attention, a quantity that is proportional to actual number of queries between mutual funds comes at the cost of being unable to capture as fully the variation in attention in the time series. The relative scaling of multiple queries employed to compare many different queries relative to the same absolute benchmark causes a coarsening of the variation over time series of the reported search volumes with respect to an individual query. In a perfect world, one would be able to obtain a panel of number of search queries posted at a given time[35], however, given the inherit discretization from 1-100 of the Google Trends data, I must work with what I have.

Google Trends can measure up to five queries at once, e.g., the NASDAQ tickers are separated within one query by up to four commas. To begin, I choose one benchmark "TICKER" and then compare the monthly frequency observations relative to all other 10,782 tickers. A good choice for the first benchmark is the fund ticker with the largest average search volume within the monthly time series relative to all other funds' "TICKER" queries for different tickers. If Google Trends did not have the censored data issue, this would be simple, but when doing this, if the

[35] The Chinese search engine Baidu can provide such data to the public. See Zhang, Shen, Zhang and Xiong, (2013). http://index.baidu.com/

search volume for "FUNDX" is not within 2 orders of magnitude of the maximum of "FUNDZ" (presumably, but not always the maximum monthly observation of the benchmark query) then it will appear as a 0 in the raw data.

Then, I make the arbitrary choice that if the average search volume is within one order of magnitude of the maximum of the 5 queries (e.g., $\overline{SV}_m^i > 10$ for fund ticker *i* of monthly frequency search volume), and claim that the search volume for these fund tickers are "well identified". If the data generated by Google Trends in the multiple comparison is not "well identified" I will toss the results and proceed to the next step. Next, I choose a new benchmark query which, based on the current benchmark, has the smallest average search volume of queries with an average search volume greater than 10 and, importantly, no monthly observation of 0. I collect the monthly frequency observations of the updated benchmark and the query generated from each fund ticker in the subsample, with the previous well identified ones saved, and iterate. Thus, I am able to move down roughly an order of magnitude in search volume in the re-querying of relative to the new benchmark. Iterate this process of selecting a new benchmark for a smaller subsample of firm tickers until there are no more benchmarks that meet the aforementioned criteria, namely having all non-zero monthly observations.

From these data, to obtain my measure of relative search volume, I utilize the ratio of successive benchmarks within the same month[36]. Denote SV_{m,b_k}^i as the search volume for stock *i* during month *m* relative to benchmark b_k, which is the kth benchmark. Denote $SV_{m,b_{k-1}}^{b_k}$ as the search volume of the kth benchmark relative to the previous benchmark. My measure of Absolute Attention of asset *i* at month *m* relative to the kth benchmark b_k is given by

$$AA_{m,b_k}^i \equiv SV_{m,b_k}^i \prod_{j-2}^{k} \frac{SV_{m,b_{j-1}}^{b_j}}{SV_{m,b_{j-1}}^{b_{j-1}}}$$

When the query corresponding to *i* and the kth benchmark is "well identified" it will be the case that:

[36] Note here dividing by 0 may be a problem, and thus this constraint in the updating process of the benchmark is binding.

$$\sum_m SV^i_{m,b_k} > \sum_m SV^{b_{k+1}}_{m,b_k}$$

And the "well identified" query will be denoted

$$AA^i_m = AA^i_{m,b_k}$$

The AA^i_m is the measure of Absolute Attention level of query corresponding to firm *i* in month *m*. This measure can be compared to $AA^{i'}_m$, where $i \neq i'$ for general queries greater than two orders of magnitude. Thus, I obtain a cross-sectional measure of Absolute Attention levels between any two assets at any two months in sample.

4) RESULTS

4.1) Absolute Attention in the Mutual Fund Industry

Armed with this measure of Absolute Attention is approximately proportional to actual search volumes between funds up to a constant (reflecting the initial choice of benchmark), I can perform aggregate analysis answering the question where within the mutual fund attention being garnered. Table 1 provides a summary of the total attention of funds across categories scaled by total Absolute Attention in the sample. Using the labels provided in the CRSP Survivor-bias-free US Mutual Fund data set, I can make frequentist inferences assuming each query into Google is statistically independent and the censoring threshold is small.

By my measure, attention is dominated in equity funds, 72.53% of the Absolute Attention within my sample period takes place in pure equity funds, while only 9.5% of Absolute Attention takes place in Bond funds. These correspond to 3003 equity funds on average per month in my sample. Attention to Equity is over-weighted as Equity funds only represent 61% of the average number of funds per month in my sample and 57% of the average number of funds per month in CRSP. Relative comparisons across fund types can also be made from this table; for example, Government (non-municipal) Bond funds are 7.2[37] times more likely to be searched for at any

[37] 7.2 = 4.32% / 0.60%

given time on Google than Corporate Bond funds. Index funds[38] are 2.5 times more likely to be command Investor Attention than non-index funds. Moreover, 95.8% of all searches are for funds which are institutional funds. Domestic Equity Institutional funds alone account for half of all investor attention of individual mutual funds. While it may appear surprising at first glance that Google searches, the proxy for "retail" investor attention as claimed by DEG 2011 and Ben-Rephael, Da, and Israelsen (2016), has such a high concentration in Institutional funds, I hypothesize that the channel that may be attracting this attention is Defined Contribution Plans and Individual Retirement Accounts. According to Cohen and Schmidt (2009) the 401(k) market is an economically large and growing channel of captive capital, account for a substantial percentage of assets and flow in the mutual fund industry. In 2015, $2.8 trillion of mutual funds were invested in 401k(s), with 59% coming from Equity funds[39]. In these specialized investment vehicles retail investors do have access into institutional funds.

For the remainder of the paper, I will focus on only Equity mutual funds which are open to new investment. Funds closed to new investment account for less than 1% of the aggregate Absolute Attention.

4.2) Performance of 12B-1 Funds in Investor Attention

I next focus on the question do Equity Mutual Funds that have 12B-1 fees achieve their goals of superior marketing relative to funds without 12B-1 fees. 12B-1 fees are 'fees paid for marketing and selling fund shares, such as compensating brokers and others who sell fund shares, and paying for advertising, the printing and mailing of prospectuses to new investors, and the printing and mailing of sales literature.'[40] Attention and brand awareness is often a goal of marketing efforts, so my Google Trends data provides insight into whether investors are reaping the rewards for the cost of these added services. This does not appear to be the case.

---

[38] Flagged in CRSP as funds that either use an index as a filter for the purchase and sale of securities or funds that seek to meet and/or exceed index.
[39] https://www.ici.org/pdf/per22-04.pdf
[40] https://www.sec.gov/fast-answers/answersmffeeshtm.html

Barber, Odean and Zheng (2005) claim that investors are learning over time to avoid funds with large load fees, but they do 'buy funds that attract their attention through advertising and distribution', e.g., the 12B-1 fee funds. My out of Barber, Odean and Zheng (2005) sample data support the claim that investors are learning to avoid fees over time, but have more recently also learned to pay less attention to funds with 12B-1 fees in addition to funds with load fees. Figure 7 splits the sample between those funds that have a 12B-1 fee indicated in CRSP and those that do not, and plots levels of Absolute Attention averaged every year in my sample. To make the measure interpretable, values of Attention are scaled to the average value of Absolute Attention in 2016 for the set of funds that have 12B-1 fees. Notably, attention is decreasing over time in the 12B-1 funds and dramatically increasing in the non-12B-1 funds. In 2016, a ticker corresponding to a fund without 12B-1 fees is 18 times more likely to be searched for than a ticker corresponding to a fund with 12B-1 fees, indicating investors are paying less attention to the funds that are spending more money on marketing services. The fact that funds with 12B-1 fees are getting less and less attention over time, and not keeping pace with their non-12B-1 fee counterparts appears suggests that either (1) these marketing costs, borne by fund shareholders, are not adding value in channels that precipitate organic attention via Google Searches, or (2) the marketing efforts of mutual funds are not focused on digital media awareness, thus my Google Search measure is not a reasonable target, which in my view is implausible given the aggregate and ubiquitous nature of Google searches as a vehicle to obtain on-demand low cost information for investors. In the DEG (2011) view, the 'retail' investors' attention, for which Google searches are proxying, are the investors that mutual fund brokers and other mutual fund marketing efforts are often targeting.

Fama and Jensen (1983) claim that the redeemable claims feature of mutual funds is a sufficient form of market governance, requiring less regulation in this market than other markets. The fact that investors are *choosing* to not pay attention to funds with 12B-1 fees indicates that they are learning from the numerous studies[41] which link fees negatively with performance.

[41] See Malkiel (1995), Carhart (1997) among many others

4.3) Attention in Equity Mutual Funds

What fund characteristics drive these two attention measures? In Table 16, I explore this question in a Panel OLS Fixed Effects predictive regressions. The measure of Absolute Attention has lots of variation, in terms of orders of magnitude of searches, just above the censored value of 0. I choose to make the modeling choice of transforming the Absolute Attention variable by $\log(AA_t + \varepsilon)$, where ε corresponds to the minimum value of Absolute Attention attained in my analysis. A log transformation is a natural transformation of highly right skewed data[42], and offsetting by the minimum value will avoid the log(0) problem in order to maintain power in observations and maximize spread of the measure. This may result in overestimating Absolute Attention levels of the funds with the lowest levels of attention e.g., at the censored value, but the shift does not affect any of the marginal effects of the coefficients in the regression analysis.

In Panel A of Table 16, I examine the features that determine Absolute Attention level under this transformation. I find that my model can explain 29% of the variation using month cross style[43] fixed effects, and a relatively parsimonious set of firm characteristics. I find that indicators on broker sold funds[44] have a significant negative relation, corresponding to a 40% reduction in search intensity, with a similar negative marginal effect for lagged 1-month returns. I find that fund size, measured by total net assets, institutional funds and daily return volatility in the prior month significantly predict positive relations with Absolute Attention. Expense ratios are highly negatively significant, while there is some evidence that older funds garner less attention than newer funds, ceteris paribus. Last fiscal year turnover appears to have a significant negative relation to predicting absolute attention, but the coefficient isn't economically large.

Panel B of Table 16 investigates what fund characteristics affect Abnormal Investor Attention, which are largely different than in Panel A. Coefficients on returns, log size, and

---

[42] For visual evidence of the benefit of log transformation here, view Figure 8, which displays the 30th and 70th cross sectional percentiles of Absolute Attention, and the corresponding log transformations coded in blue and yellow.
[43] Style refers to 4-digit CRSP Objective Code.
[44] Broker sold funds are defined as funds that have either a front load, rear load, or a 12B-1 fee.

volatility have opposite signs relative to Panel A, suggesting that changes to levels of attention are positively affected by high Sharpe ratios in relatively small funds. Expense ratios, broker sold status, momentum effects, turnover and age are measured as insignificant to changes in levels of attention. Again, institutional funds tend to be where the majority of attention resides in my sample, affecting both levels and changes to levels. The last specification includes today's level of Absolute Attention as a control predicting Abnormal Investor Attention, while admittedly highly endogenous, has a strong negative association with next period's changes in attention.

4.4) Mutual Fund Flows and Attention

As Total Net Assets is so strongly correlated with Absolute Attention, as shown in the previous section, I proceed to perform analysis on Dollar Flows with respect to levels of Absolute Attention. I compute dollar value of (net) flows for fund *i* in month t, F_t^i following Frazzini and Lamont (2008)[45] as

$$F_t^i = TNA_t^i - \left(1 + R_t^i\right)TNA_{t-1}^i - MGN_t^i$$

where TNA_t^i is the total net assets of a fund *i* in month *t*, R_t^i is the fund *i*'s return between months *t*-1 and *t*, and MGN_t^i is the increase in total net assets due to fund merges during month *t*[46]. I assume that existing investors reinvest dividends and other distributions in the fund and all inflows/outflows occur at quarter end. New funds are assumed to have inflows equal to their initial total net assets and funds that die have outflows equal to their terminal TNA.

I independently double sort returns and attention each month into portfolios based on 30/40/30 breakpoints. The monthly portfolio breakpoints in attention can be viewed in Figure 8. We can see there is some noise in the level of Absolute Attention in the early part of the sample, with higher levels of the distribution attention than towards the tail end. By October 2006 and January 2008 for the 70th and 30th (log) percentiles respective, Absolute Attention levels appear to settle down to roughly the same level for the remainder of the sample. Table 17 reports the

[45] see Zheng (1999); Sapp and Tiwari (2004)

[46] In unreported analysis, I found no significant difference in Absolute Attention levels of both the set of Target Funds nor Acquiring Funds around acquisition month.

results for next month's dollar flows on the aforementioned double sorted portfolios. I also include all funds that do not have Absolute Attention data under the 'No Attention' portfolio as, at least in an ordinal sense, they have been measured to be below the data that has crossed the Google threshold boundary. Panel A: examines the Cross Sectional Average Flow in Millions of Dollars for the average fund in each of the 3x3 portfolios+3 No Attention portfolios. There does appear to be evidence of a monotonic dollar flow to performance sensitivity, as portfolios containing the highest raw returns generate more inflows on average than other portfolios conditional on attention, and low raw returns yield average negative outflows. The No Attention portfolios appear to have average returns close to 0, while there is clearly a trend that high Attention portfolios have large absolute value of net flows than Low Attention portfolios. Panel B provides the Percentage of all net dollar flows generated by these portfolios, and Panel C describes the time series average of the number of firms within each portfolio. I detect 69% of the total amount of monthly dollar flows are coming from the High Attention Portfolios that are not in the lowest 30th percentile of returns, corresponding to just 239 out of 8,994 funds on average. Similarly, 68% of the total amount of monthly dollar flows are coming from the top 30th percentile of returns with the top 70th percentile of attention, corresponding to just 275 out of 8,994 funds on average. To get a sense of the economic 'specialness' of these portfolios.

With these non-parametric suggestive results, I next investigate to see if there exists visual evidence of the convex dollar flow to performance sensitivity in raw returns. Figure 9 reports a Penalized B-spline regression smoothing technique amongst the funds with "High Attention" (Above 70th percentile), "No Attention" as defined before, and Portfolios 1-2 from Table 17 have been grouped together as "Low Attention". Following standard practice, the dataset here has been trimmed at the top and bottom 1% levels of attention and flows, because the effects of outliers at the tails can impact the shape of the technique dramatically. In terms of dollar flows, I detect convexity only in the High Attention Funds, while No Attention funds appear to have no relationship to returns and flows. The Low Attention funds between -20% and 20% appear to have an almost linear relationship, with some non-convexity at the tails of the return distribution, even after trimming the dataset.

Table 18 employs the OLS predictive fixed effect regression framework with Dollar Flows as the dependent variable, on the sample of funds with data in Google Trends. I employ month cross Fund Style fixed effects as before, with Absolute Attention, Return, and Return squared as independent variables, the return square term is there to test convexity. In all four specifications the square term is statistically insignificant, and has the wrong sign (convexity would imply a positive sign), suggesting further evidence that the flow to performance sensitivity of mutual funds is linear in returns. I also compare Gross Return (raw return + expense ratio) with a raw return specification to discover that the log(Absolute Attention + ε[47]) measure captures more of the variation in flows in the Gross Return framework. The marginal effect, e.g., upscaling of logs, is $2.75M, which is reduced to $1.31M in a specification with more firm characteristic controls. Flows are strongly negatively predicted by firm age and expense ratio, and positively predicted by fund momentum (measured from 12 lags to 1 lag), size (mechanically) and Absolute Attention. Large, new funds, with consistent performance and low fees generate more inflows. These characteristics, save consistent performance, tend to be the same ones contributing to Absolute Attention, so there is endogenous regressors suggesting a linear model with Absolute Attention on the right-hand side may be mis-specified.

4.5) Mutual Fund Change in Market Share and Attention

For robustness, I examine Change in Market Share instead of Dollar Flows via the same Penalized B-spline method in Figure 10, following the work of Spiegel and Zhang (2013). The Market Shares are much noisier, while there is minor visual evidence that high attention funds get more increases in market share conditional on return relative to low attention funds, especially towards the middle of the return distribution. However, the Penalized B-spline appears to be fitting a lot more noise than the smooth trends in Figure 10. Change in Market share within the -20% to 20% return window appears to be roughly flat at 0 for the funds labeled "No Attention" suggestion the importance of Investor Attention in generating fund flows.

[47] In my dataset, ε is measured to be 0.000110294

Table 19 performs a Horse Race between levels (log Absolute Attention) and changes (Abnormal Investor Attention) against both Dollar Flows and Changes in Market Share. In a full panel setting the squared returns coefficient continues to be negative. The squared raw returns are negative and significant in the Change in Market Share regression, indicating there is concavity in the tails of the return distribution, not convexity, whereby investors expect some mean reversion from extreme performance. Including an attention factor by itself appears to improve predictability in sample. In the full specification for Dollar Flows and Change in Market Share, I include both attention measures, and two new controls, the volatility of the daily series of attention and the percent of the month with searches below Google's threshold. Both factors load negatively on both dependent variables. In this specification AIA_t beats out $\log(AA_t + \varepsilon)$, which has a negative, but statistically insignificantly different from zero sign in the Changes in Market Share model. Changes in attention are more informative than levels of attention with respect to fund flows or changes in market share.

5) Conclusions

Using Google Trends search volume to inform investor attention is an exciting and developing area of research. I have documented a novel process to construct relative comparisons amongst large quantities of search queries with Google Trends, a service which appears at face value to only be able to provide two orders of magnitude of variation. Applying this process for mutual funds, I have documented the first paper, to my knowledge, that empirically estimates relative investor attention levels across fund types and characteristics. Using Absolute Attention in conjunction with Abnormal Investor Attention provides a clearer picture in understanding the process of how attention is changing in the cross section and the time series. I've also documented evidence that 12B-1 fees are not accomplishing their stated purpose of value-added marketing for mutual funds in the Internet age, suggesting that 12B-1 plans or marketing strategies employed by 12B-1 funds need to be amended or removed. With the ease of getting information from Internet search, and more investors growing up technologically savvy, there appears to be a trend towards learning how to avoid obfuscated fees which are associated with negative performance. The market disciplines mutual fund industry, once the information required to understand can be accessed easily.

Table 15: Attention by Mutual Fund Class (2-digit CRSP Objective Code)

	Percent Absolute Attention					Relative Attn		Funds/Month	
Fund Type	Total	Index	Non Index	Instl	Non Instl	Index/ Non Index	Instl/ Non Instl	Avg From Sample	Avg From CRSP
Domestic Equity	52.99%	38.85%	14.15%	50.00%	2.57%	2.7	19.4	2247.6	7290.6
Foreign Equity	19.54%	12.91%	6.63%	21.70%	0.24%	1.9	90.4	756.1	2367.0
Fixed Income (unspecified)	3.10%	1.68%	1.42%	3.10%	0.29%	1.2	10.6	393.4	1431.4
Corporate Bonds	0.60%	0.34%	0.26%	0.60%	0.02%	1.3	32.3	78.7	308.4
Foreign Bonds	2.09%	0.60%	1.48%	2.19%	0.04%	0.4	58.6	121.1	413.5
Govt. Bonds	4.32%	4.09%	0.23%	4.10%	0.06%	17.7	64.0	125.6	413.5
Money Market	0.17%		0.17%	0.17%	0.05%		3.5	238.1	1158.4
Municipal Bonds	2.23%	1.34%	0.89%	2.28%	0.47%	1.5	4.8	355.9	1565.6
Mixed Class (unspecified)	14.56%	11.78%	2.78%	11.38%	0.46%	4.2	24.8	440.1	1558.3
Target Funds	0.00%		0.00%	0.00%	0.00%		0.1	6.3	13.2
Other (unspecified)	0.33%	0.05%	0.27%	0.19%	0.00%	0.2	326.9	52.2	233.6
Currency	0.02%	0.00%	0.02%	0.02%	0.00%	0.0	69214	5.0	19.3
Mortgage Backed	0.04%	0.04%	0.00%	0.04%	0.01%	8.8	7.5	60.0	236.3
Total	100.0%	71.68%	28.32%	95.79%	4.21%	2.5	22.7	4880.2	17009.2

Table displays a summary of the Absolute Attention measure across Mutual Fund Types, determined by 2-digit CRSP Objective Code. Indicators for Institutional Funds and Index funds (any of the three index indicators) are obtained from CRSP. Interpreting this table, for Google Searches for Mutual Fund NASDAQ tickers, I estimate 53% of searches come from Domestic Equity Funds, 71.7% come from Index funds, and 95.8% come from Institutional Funds. A search for a Domestic Equity Fund is 2.7 times more likely to be an Index fund versus a non-index fund and 19.4 times more likely to be an institutional Fund versus a non-institutional fund. The Sample consists of Mutual Fund tickers that exceeded the threshold of attention level according to Google Trends from the sample period of Jan 2004-March 2017.

Table 16: Determinants of Equity Mutual Fund Attention

Panel A: Absolute Attention

Dependent Var	ln(AA^i_{t+1}+ϵ)		Marg Effect	ln(AA^i_{t+1}+ϵ)		Marg Effect
	β	t-stat	e^β	β	t-stat	e^β
Raw Return	-0.504	-3.76	0.6	-0.623	-4.89	0.5
ln(Total Net Assets)	0.153	74.96		0.127	60.16	
ln(Fund Age)	-0.016	-2.56		0.004	2.68	
Momentum$_{t-12,t-1}$	-0.010	-0.70	1.0	-0.015	-1.05	1.0
Index	3.403	190.50	30.1	3.316	184.05	27.5
Broker Sold	-0.581	-55.87	0.6	-0.523	-49.78	0.6
Institutional Fund	0.799	72.22	2.2	0.725	64.27	2.1
Daily Return Volatility	7.915	10.55	2737.2	9.286	12.37	10787.2
Last FY Turnover	-0.003	-9.32	1.0	-0.003	-9.91	1.0
Expense Ratio				-24.685	-41.26	0.0
Observations	422235			420487		
R-squared	28.83%			29.14%		
Root MSE	2.674			2.680		
Month x Style FE	Y			Y		

Panel B: Abnormal Investor Attention

Dependent Var	AIA_{t+1}		AIA_{t+1}		AIA_{t+1}	
	β	t-stat	β	t-stat	β	t-stat
Raw Return$_t$	0.103	3.78	0.106	3.86	0.077	2.83
ln(Total Net Assets)	-0.003	-7.19	-0.003	-6.84	0.001	2.95
ln(Fund Age)	-0.001	-0.75	-0.002	-1.17	-0.001	-1.02
Momentum$_{t-12,t-1}$	0.000	-0.16	0.000	-0.04	-0.001	-0.22
Index	0.004	1.21	0.004	1.21	0.118	31.16
Broker Sold	0.001	0.57	0.001	0.29	-0.017	-8.08
Institutional Fund	0.010	4.33	0.010	4.23	0.034	15.02
Daily Return Volatility	-0.363	-2.39	-0.404	-2.63	-0.050	-0.33
Last FY Turnover			0.000	0.17	0.000	-1.54
Expense Ratio			0.046	0.36	-0.847	-6.80
ln(AA^i_t+ϵ)					-0.034	-109.39
Observations	414699		417922		417922	
R-squared	2.40%		2.40%		5.15%	
Root MSE	0.547		0.547		0.539	
Month x Style FE	Y		Y		Y	

Table displays predictive OLS regressions of attention. Dependent variable occurs at month $t+1$. All independent variables occur at month t unless otherwise stated. In Panel A, dependent variable is the natural log of (Absolute Attention + ϵ), where ϵ corresponds to the minimum non-zero Absolute Attention measure observed, e.g., the smallest point above the Google Trends reporting threshold in sample. In Panel B, dependent variable corresponds to DEG (2011) measure of Abnormal Investor Attention, given by month log of 1+$t+1$ search volume index subtracted from the log of 1 + average previous two months search volume, generated from single ticker query. Daily Return Volatility is the standard deviation of daily returns in month t. Last FY Turnover indicates the turnover percentage of the last fiscal year, ranging from 0 to 1. Momentum is measured as the product of the raw returns in the last 11 months before the current month. Broker Sold is an indicator for the presence of a Front Load, Rear Load, or 12B-1 Fee. Index is an indicator of an index fund. Style refers to 4-digit CRSP Objective Code. Institutional Fund is an indicator for Institutional Fund coded as 'Y' in CRSP. Marg Effect indicates the Marginal Effect of a unit change in an independent variable. Fund Age is measured in terms of years since inception. Total Net Assets is measured in ($M).

Table 17: Equity Mutual Fund Flows - Double Sort on Returns and Absolute Attention

Panel A: Cross Sectional Average Flows ($M) Per Fund Per Month within Portfolio

		High Raw Returns		Low Raw Returns
		3	2	1
High Attention	3	17.94	10.34	-4.65
	2	7.39	-2.03	-3.64
Low Attention	1	4.86	-0.32	-0.73
No Attention	0	0.66	0.15	-0.26

Panel B: Percentage of Cross Sectional Average Net Flows Attributed to Each Portfolio

		High Raw Returns		Low Raw Returns	Total
		3	2	1	
High Attention	3	48.07%	20.98%	-13.09%	55.96%
	2	20.32%	-5.64%	-9.70%	4.98%
Low Attention	1	15.34%	-1.16%	-2.14%	12.04%
No Attention	0	29.63%	9.18%	-11.78%	27.03%
	Total	113.36%	23.36%	-36.72%	100.00%

Panel C: Average Number of Funds per Portfolio per Month

		High Raw Returns		Low Raw Returns	Total
		3	2	1	
High Attention	3	136	103	145	383
	2	139	141	135	414
Low Attention	1	160	185	149	494
No Attention	0	2263	3168	2271	7702
	Total	2689	3597	2700	8994

Sample consists of Mutual Funds from Jan 2004-March 2017. Funds are independently sorted on raw returns and Absolute Attention into 9 portfolios. If there is missing data for Absolute Attention, e.g., the ticker query did not pass the threshold from Google Trends, then all remaining funds are allocated to the 'No Attention' portfolio, portfolio 0. Portfolio 1 corresponds to the bottom 30% of Absolute Attention (raw returns); Portfolio 2 corresponds to the middle 40% of Absolute Attention (raw returns); Portfolio 3 corresponds to the top 30% of Absolute Attention (raw returns).

Table 18: Dependent Variable - Equity Mutual Fund Flows ($M)

Regression	1		2		3		4	
	β	t-stat	β	t-stat	β	t-stat	β	t-stat
Raw Return	47.431	4.50	75.702	5.78	75.129	5.72		
$(\text{Raw Return})^2$	-2.550	-1.51	-3.525	-1.71	-3.472	-1.68		
ln(Absolute Attention + ϵ)	0.863	9.90	0.317	3.21	0.267	2.67	1.010	10.37
ln(Total Net Assets)			2.903	20.69	2.812	18.94		
ln(Fund Age)			-14.095	-35.69	-13.866	-34.26		
Daily Return Volatility			-112.040	-1.96	-112.122	-1.96		
Last FY Turnover			-0.035	-1.68	-0.035	-1.70		
$\text{Momentum}_{t-12,t-1}$			4.438	4.53	4.404	4.49		
Expense Ratio					-118.378	-3.03		
Gross Return							37.858	3.02
$(\text{Gross Return})^2$							-1.842	-1.01
Observations	445371		412967		408356		407213	
R-squared	1.67%		2.10%		2.11%		1.78%	
Root MSE	176.26		180.75		181.5212		182.64	
Month x Style Fixed Effects	Y		Y		Y		Y	
Marginal Effect of Absolute Attention Flows ($M)	2.37		1.37		1.31		2.75	

Table reports one month ahead predictive OLS regressions for dependent variables of Dollar Flows at time $t+1$ in millions of dollars (Frazzini and Lamont (2008)), to test the convexity (squared raw returns) in flow to performance sensitivity using log(AA_t + ϵ) as a control, where ϵ corresponds to the minimum non-zero Absolute Attention measure observed, e.g., the smallest point above the Google Trends reporting threshold in sample. Fund Age is measured in years since fund inception. Daily Return Volatility is the standard deviation of daily returns in month t. Last FY Turnover indicates the turnover percentage of the last fiscal year, ranging from 0 to 1. Total Net Assets is measured in Millions. Gross Return is defined as Raw Return plus Expense Ratio, e.g., the return investors see before expenses. Momentum is measured as the product of the raw returns in the last 11 months before the current month. Expense ratio is the ratio of fund expenses as defined in CRSP. Style refers to 4-digit CRSP Objective Code. Marginal effects for the coefficient on Absolute Attention are reported for interpretation of the log-scaled variable.

Table 19: Horse Race - Absolute Attention vs. Abnormal Attention

Dependent Var	Flow ($M)				Change in Market Share (bp)			
Raw Return	31.949	47.431	50.543	75.137	1.069	1.564	1.635	2.295
	(6.54)	(4.50)	(4.68)	(5.81)	(9.68)	(6.53)	(6.67)	(7.66)
$(\text{Raw Return})^2$	-1.324	-2.550	-2.755	-3.408	-0.088	-0.127	-0.132	-0.144
	(-1.45)	(-1.51)	(-1.61)	(-1.67)	(-4.22)	(-3.29)	(-3.39)	(-3.05)
$\ln(AA_t + \epsilon)$		0.863		0.058		0.011		-0.004
		(9.90)		(0.52)		(5.69)		(-1.71)
AIA_t			1.956	1.763			0.046	0.048
			(3.91)	(3.29)			(4.13)	(3.85)
ln(Fund Age)				-13.800				-0.273
				(-34.45)				(-29.47)
ln(Total Net Assets)				2.728				0.026
				(18.51)				(7.57)
Daily Return Volatility				-115.901				-3.282
				(-2.05)				(-2.51)
Last FY Turnover				-0.043				-0.001
				(-2.11)				(-1.5)
$\text{Momentum}_{t-12,t-1}$				4.428				0.036
				(4.57)				(1.59)
Expense Ratio				-150.316				-3.917
				(-3.69)				(-4.16)
% of month with "No Attention"				-3.989				-0.152
				(-2.19)				(-3.61)
Daily SVI Volatility				-0.134				-0.004
				(-3.06)				(-4.23)
Observations	1436752	445371	438055	415451	1481416	457695	450328	415451
R-squared	0.70%	1.67%	1.67%	2.12%	1.38%	3.01%	3.04%	3.48%
Root MSE	101.447	176.26	177.5323	179.9683	0.023	0.040	0.041	0.042
Month x Style FE	Y	Y	Y	Y	Y	Y	Y	Y

Table reports one month ahead predictive OLS regressions for dependent variables of Dollar Flows at time $t+1$ in millions of dollars (Frazzini and Lamont (2008)) and changes in Market Share at time $t+1$ in basis points (Spiegel and Zhang (2013)). Unless otherwise stated, all independent variables occur in month t. The $\ln(AA_t + \epsilon)$ variable refers to natural log of (Absolute Attention + ϵ), where ϵ corresponds to the minimum non-zero Absolute Attention measure observed, e.g., the smallest point above the Google Trends reporting threshold in sample. AIA_t corresponds to DEG (2011)'s measure of Abnormal Investor Attention, given by month log of 1+$t+1$ search volume index subtracted from the log of 1 + average previous two months search volume, generated from single ticker query. Fund Age is measured in terms of years since inception. Fund Size is measured in ($M). Daily Return Volatility is the standard deviation of daily returns in month t. Last FY Turnover indicates the turnover percentage of the last fiscal year, ranging from 0 to 1. Momentum is measured as the product of the raw returns in the last 11 months before the current month. Broker Sold is an indicator for the presence of a Front Load, Rear Load, or 12B-1 Fee. Index is an indicator of an index fund. Style refers to 4-digit CRSP Objective Code. Institutional Fund is an indicator for Institutional Fund coded as 'Y' in CRSP. Daily SVI Volatility is refers to the standard deviation of the Google Trends search volume index (SVI) for a single ticker query measured at the daily frequency. Percent of month with "No Attention" refers to the number of days within the month for a single ticker query measured at the daily frequency where Google Trends reported a 0 for that day. Estimated coefficients in Change in Market Share regressions are multiplied by 100.

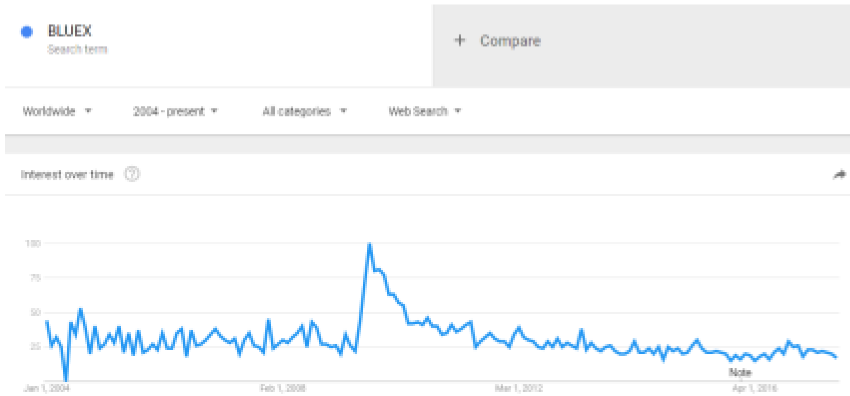

Figure 3: Google Trends Data for AMG Managers Brandywine Blue Fund Class I 'BLUEX', illustrating yields maximal variation across the time series.

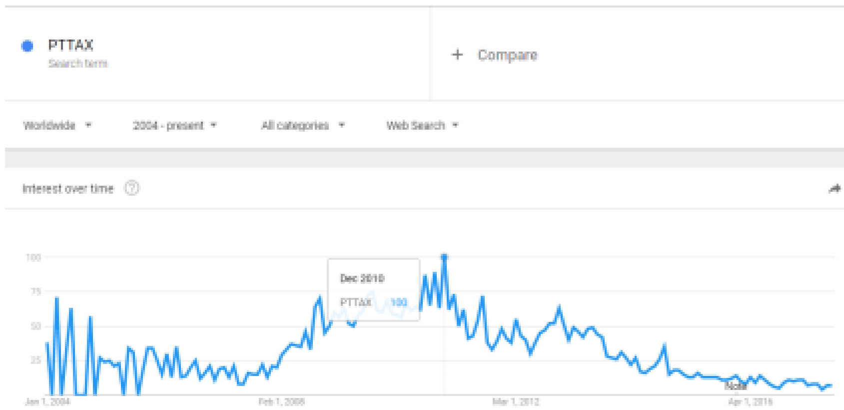

Figure 4: PIMCO Total Return Fund Class A 'PTTAX' alone, yields maximal variation across the time series.

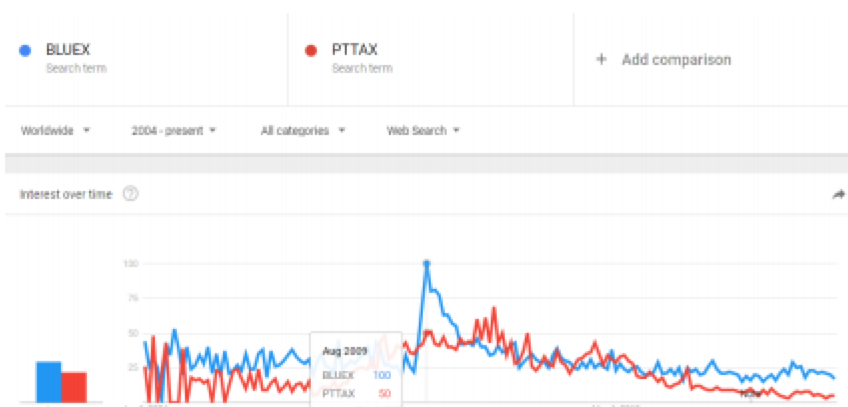

Figure 5: PIMCO Total Return Fund Class A 'PTTAX', as compared to BLUEX

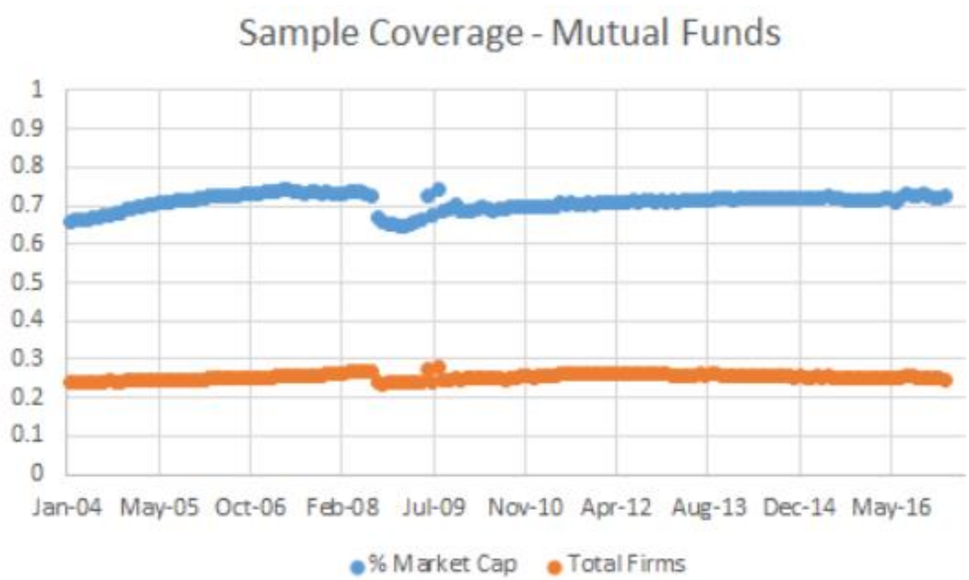

Figure 6: This Figure displays the sample coverage that the Google Trends data was able to match the full sample of the CRSP Suvivor-Bias Free Mutual Fund Database every month from Jan 2004-March 2017

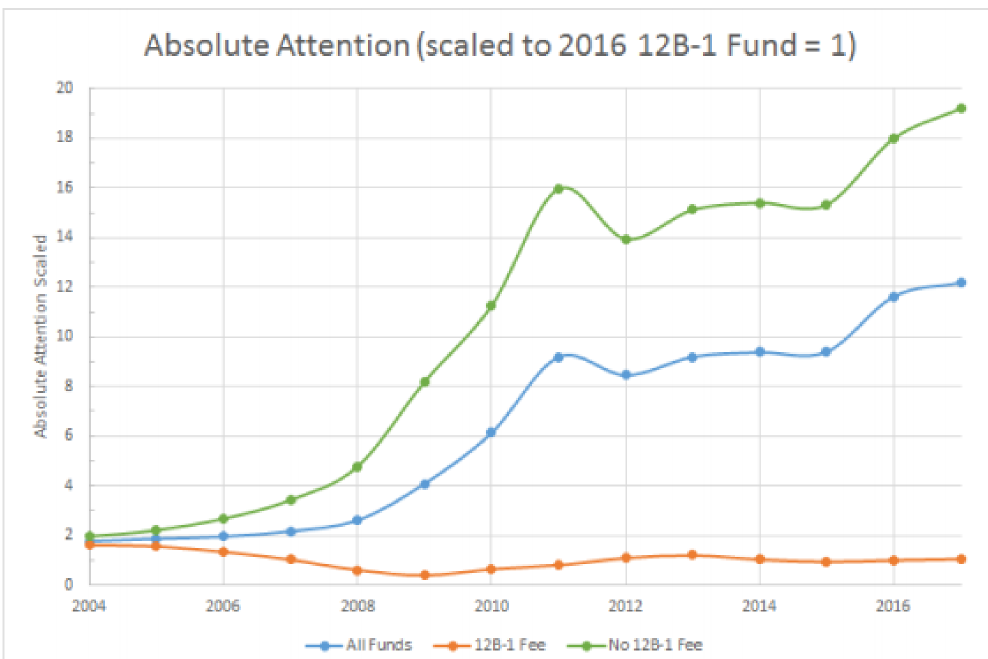

Figure 7: Within the sample of Equity Mutual Funds that have monthly search data exceeding the threshold from Google Trends for at least one month from January 2004-March 2017, this figure displays a scaled value of Absolute Attention relative to setting the value obtained for 2016 12B-1 Fee funds = 1. In 2016, a ticker from a fund without a 12B-1 fee is 18 times more likely to be searched for than a one from a fund with a 12B-1 fee. The blue line examines all funds in sample.

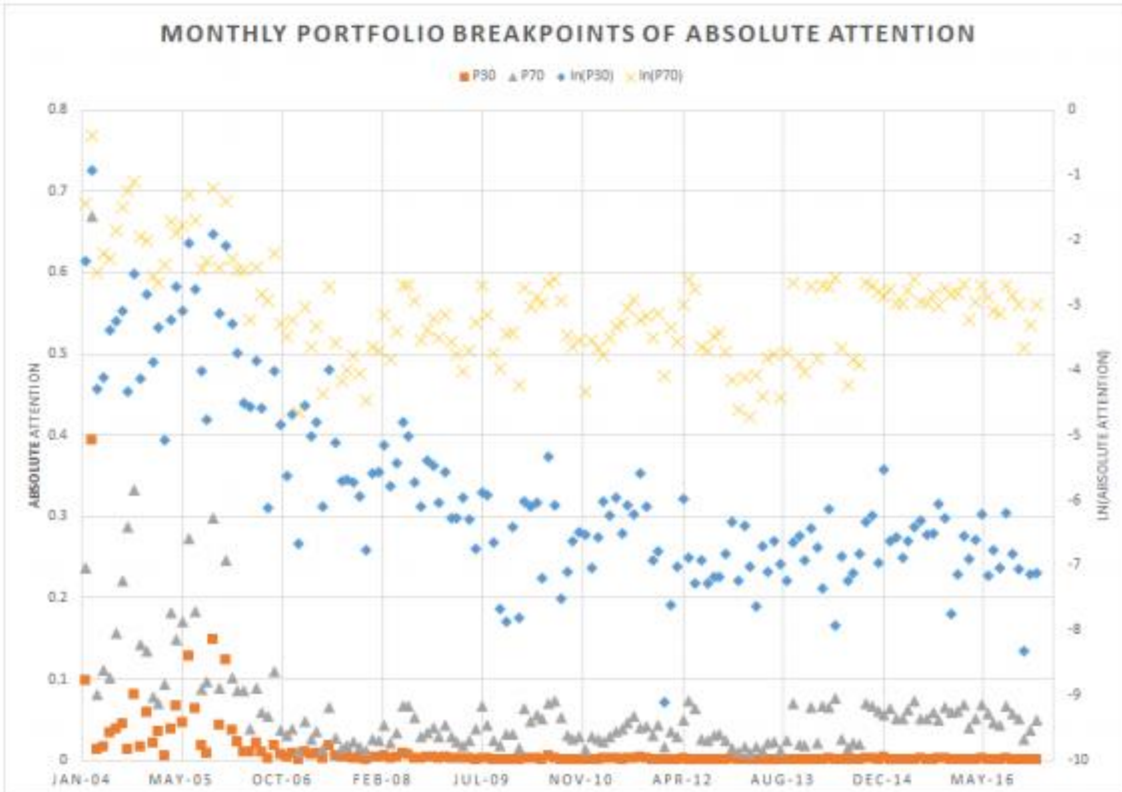

Figure 8: Displays the 30th and 70th percentiles of Absolute Attention per month used to determine the portfolio breakpoints in Table 3. The Absolute Attention measure is also displayed in logs to demonstrate the variation around the 0 absorbing barrier.

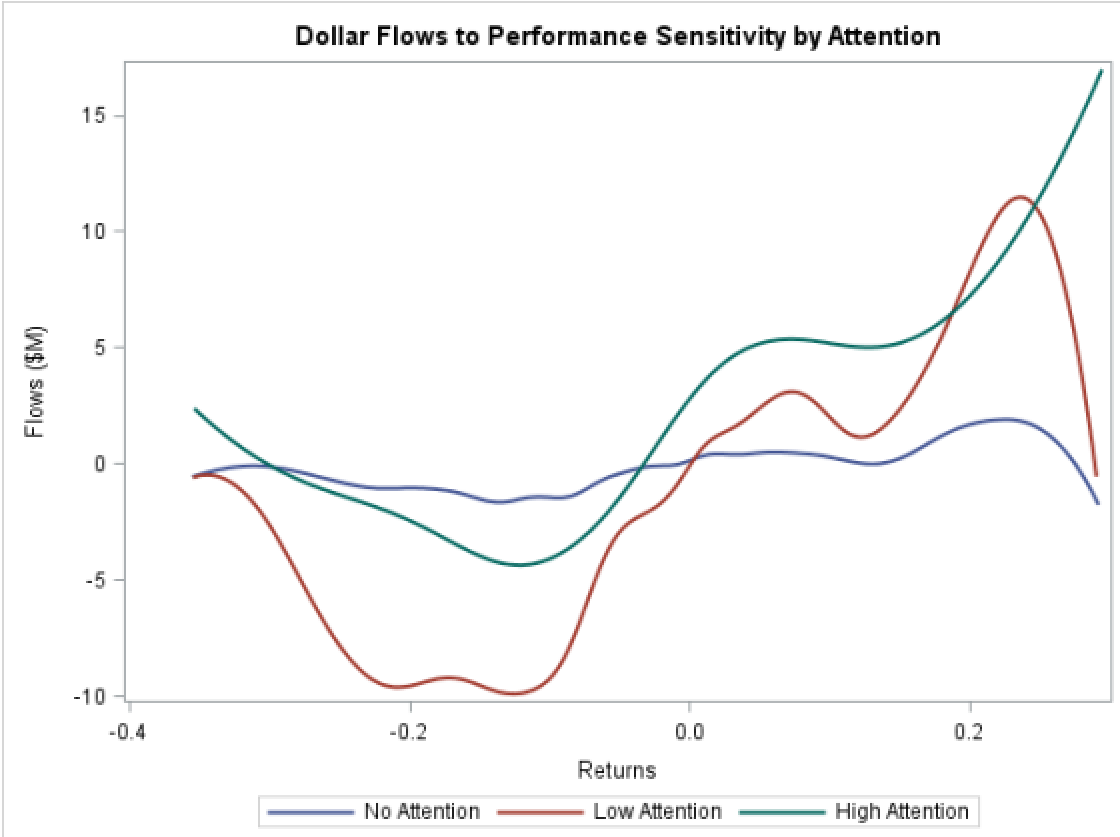

Figure 9: Displays the Dollar Flow ($M) to Performance (raw monthly returns) sensitivity via a Penalized B-spline regression smoothing technique. The "High Attention" funds are the funds that exceeded the 70th percentile in Absolute Attention. The "High Attention" funds are the funds that were below the 70th percentile in Absolute Attention, but had enough attention to exceed the Google Trends threshold for data collection. The "No Attention" funds were funds that did not exceed the Google Trends threshold level. Sample consists of Equity Mutual Funds from Jan 2004-March 2017. Following the standard practice, I have trimmed the top and bottom 1% of the data.

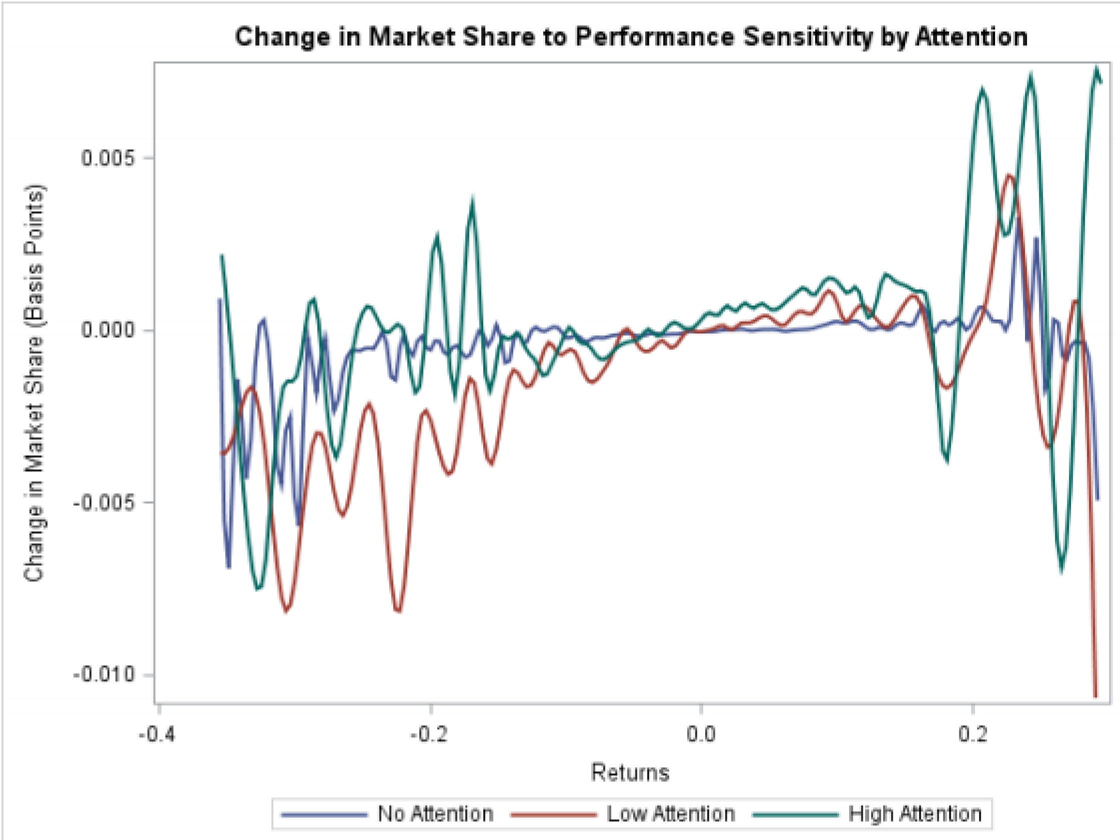

Figure 10: Change in Market Share (basis points) as calculated by Spiefel and Zhang (2013) to Performance (raw monthly returns) sensitivity via a Penalized B-spline regression smoothing technique. The "High Attention" funds are the funds that exceeded the 70th percentile in Absolute Attention. The "High Attention" funds are the funds that were below the 70th percentile in Absolute Attention, but had enough attention to exceed the Google Trends threshold for data collection. The "No Attention" funds were funds that did not exceed the Google Trends threshold level. Sample consists of Equity Mutual Funds from Jan 2004-March 2017. Following the standard practice, I have trimmed the top and bottom 1% of the data.

REFERENCES

Amihud Y., and Goyenko R. 2013. Mutual fund R2 as predictor of performance. Review of Financial Studies 26:667–94.

Barber, B., Odean, T., 2008. All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors. Review of Financial Studies, 21 (2), 785-818.

Barber, B., Odean, T., Zheng, L., 2005. Out of Sight, Out of Mind: The Effects of Expenses on Mutual Fund Flows. Journal of Business, 78, No. 6, 2095-2119.

Bauguess, S. and Cooney, J. and Hanley, K. Investor Demand for Information in Newly Issued Securities (June 28, 2018). Available at SSRN: https://ssrn.com/abstract=2379056 or http://dx.doi.org/10.2139/ssrn.2379056

Ben-Rephael, A., Da, Z. and Israelsen, R., 2017. It Depends on Where You Search: Institutional Investor Attention and Under-Reaction to News. Review of Financial Studies, 30 (9), 3009-3047.

Berk, J. and Green, R. (2004). Mutual Fund Flows and Performance in Rational Markets. Journal of Political Economy. 112. 1269-1295. 10.2139/ssrn.383061.

Cao, S. and Du, K., Yang, B., and Zhang, L., Copycat Skills and Disclosure Costs: Evidence from Peer Companies' Digital Footprints (November 8, 2018). Available at SSRN: https://ssrn.com/abstract=3280744 or http://dx.doi.org/10.2139/ssrn.3280744

Carhart, M., On Persistence in Mutual Fund Performance. 1997. Journal of Finance, 52, No. 1 57-82.

Chen, Huaizhi and Cohen, Lauren and Gurun, Umit and Lou, Dong and Malloy, Christopher J., IQ from IP: Simplifying Search in Portfolio Choice. 8th Miami Behavioral Finance Conference 2017. Available at SSRN: https://ssrn.com/abstract=3037282 or http://dx.doi.org/10.2139/ssrn.3037282

Chen, Y. and Kelly, B. and Wu, W., Sophisticated Investors and Market Efficiency: Evidence from a Natural Experiment. Journal of Financial Economics, Volume 138, Issue 2, November 2020, Pages 316-341

Crane, A., Crotty, K., and Umar, T., Hedge Funds and Public Information Acquisition (April 23, 2019). Available at SSRN: https://ssrn.com/abstract=3127825 or http://dx.doi.org/10.2139/ssrn.3127825

Da, Z., Engelberg, J., Gao, P., 2011. In Search of Attention. Journal of Finance, 66 (5), 1461-1499.

Da, Z., Engelberg, J., Gao, P., 2015. The Sum of All FEARS Investor Sentiment and Asset Prices. Review of Financial Studies, 28 (1) 1-32.

Del Guercio, D., and Tkac, P., 2008. Star Power: The Effect of Morningstar Ratings on Mutual Fund Flow. Journal of Financial and Quantitative Analysis, 43, No. 4, 907-936.

Doshi, H., Elkamhi, R., and Simutin, M., Managerial Activeness and Mutual Fund Performance, The Review of Asset Pricing Studies, Volume 5, Issue 2, December 2015, Pages 156–184, https://doi.org/10.1093/rapstu/rav005

Drake, M., Roulstone, D., Thornock, J. (2014) The Determinants and Consequences of Information Acquisition via EDGAR, Contemporary Accounting Research. 32:3, 1128-1161.

Drake, M. and Johnson, B., Roulstone, D. and Thornock, J., Is There Information Content in Information Acquisition? (March 1, 2019). Available at SSRN: https://ssrn.com/abstract=3356055 or http://dx.doi.org/10.2139/ssrn.3356055

Dyer, T., 2017, Does public information acquisition level the playing field or widen the gap? An analysis of local and non-local investors, Working Paper University of North Carolina.

Evans, R., Sun, Y., 2020. Model or Stars: The Role of Asset Pricing Models and Heuristics in Investor Risk Adjustment. The Review of Financial Studies, hhaa043, https://doi.org/10.1093/rfs/hhaa043

Fama, E., Jensen, M., Agency Problems and Residual Claims. The Journal of Law \& Economics, 26, No. 2, 329-349.

Fang, L., Peress, J., 2009. Media Coverage and the Cross-section of Stock Returns. Journal of Finance, 64 (5) 2023-2052.

Fisher, A., Martineau, C., Sheng, J., (2016) Media Attention, Macroeconomic Fundamentals, and the Stock Market. Available at SSRN: https://ssrn.com/abstract=2703978

Frazzini, A., Lamont, O., 2008. Dumb money: Mutual fund flows and the cross-section of stock returns. Journal of Financial Economics, 88, 299-322.

Gibbons, B., Iliev, P., and Kalodimos, J. Analyst Information Acquisition via EDGAR (January 2, 2019). Available at SSRN: https://ssrn.com/abstract=3112761 or http://dx.doi.org/10.2139/ssrn.3112761

Grossman, S., and Stiglitz, J. “On the Impossibility of Informationally Efficient Markets.” The American Economic Review, vol. 70, no. 3, 1980, pp. 393–408. JSTOR, www.jstor.org/stable/1805228.

Grullon, G., Kanatas, G., Weston, J., 2004. Advertising, Breadth of Ownership, and Liquidity. Review of Financial Studies, 17 (2), 439-461.

Hirshleifer, D., Teoh, S. H., 2003. Limited attention, information disclosure, and financial reporting. Journal of Accounting and Economics, 36. 337-386.

Hirshleifer, D., Lim, S., and Teoh, S., 2009. Driven to Distraction: Extraneous Events and Underreaction to Earnings News, The Journal of Finance, Vol 64, No. 5 pp 2298-2325.

Huberman, G., Regev, T., 2001. Contagious Speculation and a Cure for Cancer: A Nonevent that Made Stock Prices Soar. Journal of Finance, 56 (1), 387-396.

Iliev, P., Kalodimos, J., and Lowry, M., Investors' Attention to Corporate Governance (March 6, 2019). 9th Miami Behavioral Finance Conference 2018. Available at SSRN: https://ssrn.com/abstract=3162407 or http://dx.doi.org/10.2139/ssrn.3162407

Jain, P., Wu, J., Truth in Mutual Fund Advertising: Evidence on Future Performance and Fund Flows. 2000. Journal of Finance, 55, No. 2, 937-958.

Kacperczyk M., Sialm C., and Zheng L. 2008. Unobserved actions of mutual funds. Review of Financial Studies 21:2379–416.

Lee, C., Ma, P., and Wang, C.. "Search-Based Peer Firms: Aggregating Investor Perceptions Through Internet Co-Searches." Journal of Financial Economics 116, no. 2 (May 2015):410–431.

Li, F. and Sun, C. Information acquisition and expected returns: Evidence from EDGAR search traffic. (2017). 1-57. Research Collection Lee Kong Chian School Of Business. Available at: https://ink.library.smu.edu.sg/lkcsb_research/5322

Loughran, T., McDonald, B. (2014) Information Decay and Financial Disclosures, Working Paper https://gsm.ucdavis.edu/sites/main/files/file-attachments/03_information_decay_and_financial_disclosures.pdf

Loughran, T., McDonald, B. (2017) The Use of EDGAR Filings by Investors, Journal of Behavioral Finance, 18:2, 231-248.

Malkiel, B., 1995. Returns from Investing in Equity Mutual Funds 1971 to 1991. Journal of Finance, 50, No. 2 549-572.

Merton, R., 1987. A Simple Model of Capital Market Equilibrium with Incomplete Information. Journal of Finance, 42, 483-510.

Preis, T., Reith, D., Stanley, H. E., 2010. Complex dynamics of our economic life on different scales: insights from search engine query data. Philosophical Transactions of the Royal Society A 368, 5707-5719.

Preis, T., Moat, H., Stanley, H. E., 2013. Quantifying Trading Behavior in Financial Markets Using Google Trends. Scientific Reports 3, 1684. ISSN 2045-2322

Ryans, J., Using the EDGAR Log File Data Set (February 8, 2017). Available at SSRN: https://ssrn.com/abstract=2913612 or http://dx.doi.org/10.2139/ssrn.2913612

Sapp, T., Tiwari, A., 2004. Does stock return momentum explain the smart money effect? Journal of Finance 59, 2605-2622.

Sensoy, B., 2009. Performance evaluation and self-designated benchmark indexes in the mutual fund industry, Journal of Financial Economics 92, No. 1, 25-39.

Sicherman, N., Loewenstein, G., Seppi, D., Utkus, S., 2016. Financial Attention. Review of Financial Studies, 29, (4), 863-897.

Sirri, E. and Tufano, P., 1998. Costly search and mutual fund flows, Journal of Finance 53, 1589-1622.

Solomon, D., Soltes, E., and Sosyura, D., 2014. Winners in the spotlight: Media coverage of fund holdings as a driver of flows. Journal of Financial Economics 113, 53-72.

Spiegel, M. and Zhang, H., 2010. Mutual Fund Risk and Market Share Adjusted Fund Flows. Journal of Financial Economics, Vol. 108, No. 2.

Tetlock, P., 2007. Giving Content to Investor Sentiment: The Role of Media in the Stock Market. Journal of Finance 62 (3), 1139-1168.

Tetlock, P., 2011. Does Public Financial News Resolve Asymmetric Information? Review of Financial Studies 23, 3520-3557.

Tetlock, P., Saar-Tsechansky, M., Macskassy, S., 2008. More Than Words: Quantifying Language to Measure Firms' Fundamentals. Journal of Finance 63 (3) 1437-1467.

Wang, P., Demand for Information and Stock Returns: Evidence from EDGAR (November 13, 2019). Available at SSRN: https://ssrn.com/abstract=3348513 or http://dx.doi.org/10.2139/ssrn.3348513

Zhang, W., Shen, D., Zhang, Y., Xiong, X., 2013. Open source information, investor attention, and asset pricing. Economic Modeling, 33, 613-619.

Zheng, L., 1999. Is money smart? A student of mutual fund investors' selection ability. Journal of Finance 54, 901-933.